c search, using sw model

il 7, 2004, 17:30:20 ; Search time 4482.16 Seconds
(without alignments)
11294.703 Million cell updates/sec

09-245-198A-1
8
gtgctgagcctggcctgg........ataaatcatgattctcctc 1168

NTITY_NUC
op 10.0 , Gapext 1.0

0272 seqs, 2167151€995 residues

s satisfying chosen parameters: 6940544

th: 0
th: 2000000000

nimum Match 0%
ximum Match 100%
sting first 45 summaries

Copyright GenCore version 5.1.6
opyright (c) 1993 - 2004 Compugen Ltd.

score greater than or equal to the score of the result being pri
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1168	100.0	1168	6	BD062757	BD062757 A t
2	1168	100.0	1239	10	AF030100	AF030100 Mu
c 3	711	60.9	203083	2	AC069459	AC069459 Mus
c 4	711	60.9	234182	10	AL603707	AL603707 Mc
5	628.6	53.8	1353	6	AX201324	AX201324 Sec
6	628.6	53.8	1353	9	AY358870	AY358870 Hon
7	628.6	53.8	1368	6	AF055872	AF055872 Hon
8	628.6	53.8	1421	6	BD090952	BD090952 Apc
9	624	53.4	1306	9	AF030099	AF030099 Hon
10	614.6	52.6	1373	6	BD062758	BD062758 A t
11	597.8	51.2	1236	6	AR140407	AR140407 Sec
12	597.8	51.2	1236	6	BD057124	BD057124 Men
c 13	566.6	48.5	130254	2	AC136195	AC136195 Rat
14	566.6	48.5	165316	2	AC119115	AC119115 Rat
c 15	566.6	48.5	223877	2	AC098923	AC098923 Rat
c 16	566.6	48.5	225077	2	AC136563	AC136563 Rat
17	498.8	42.7	898	6	AX180714	AX180714 Sec
c 18	428	36.6	234801	2	AC118309	AC118309 Rat
19	409.4	35.1	1642	2	BC019047	BC019047 Hon
20	328.4	28.1	1816	9	AY081051	AY081051 Hon
21	304	26.0	218485	2	AC127470	AC127470 Par
c 22	303.8	26.0	60268	9	AC016876	AC016876 Hon
23	278.2	23.8	148855	2	AC126921	AC126921 Bos
24	261	22.3	149736	2	AC126239	AC126239 Fel
25	237.4	20.3	180222	2	AC130192	AC130192 Sus
26	212	18.2	176258	2	AC126925	AC126925 Car
27	111.6	9.6	212093	2	AC126237	AC126237 Car
28	88.4	7.6	7218	2	I66494	I66494 Seque
29	69.8	6.0	195	6	AX379024	AX379024 Sec
30	56.2	4.8	125020	9	AF429315	AF429315 Hon
c 31	55.8	4.8	228056	2	AC144061	AC144061 Mac
c 32	55.8	4.8	303091	2	AC084799	AC084799 Mus
33	55.2	4.7	115145	10	AL807240	AL807240 Mc
c 34	55	4.7	185822	2	AC073554	AC073554 Hon
35	54.8	4.7	100511	2	AC010774	AC010774 Hon
36	54.4	4.7	87120	2	AC012225	AC012225 Hon
37	54.2	4.6	191415	10	AC096625	AC096625 Mt
38	54.2	4.6	298166	2	AC087563	AC087563 Hon
c 39	53.6	4.6	205350	2	AC078946	AC078946 Mus
40	53.2	4.6	49430	2	AC100434	AC100434 Mus
c 41	53.2	4.6	189897	9	AC084064	AC084064 Hon
42	53	4.5	55061	2	AC091597	AC091597 Mus
43	53	4.5	161575	2	AC141041	AC141041 Rat
c 44	53	4.5	187252	10	AL607109	AL607109 Mc
45	52.8	4.5	84514	10	AL627264	AL627264 Mc

nEmbl:*
gb_ba:*
gb_htg:*
gb_in:*
gb_om:*
gb_ov:*
gb_pat:*
gb_ph:*
gb_pl:*
gb_pr:*
gb_ro:*
gb_sts:*
gb_sy:*
gb_un:*
gb_vi:*
em_ba:*
em_fun:*
em_hum:*
em_in:*
em_mu:*
em_om:*
em_or:*
em_ov:*
em_pat:*
em_ph:*
em_pl:*
em_ro:*
em_sts:*
em_un:*
em_vi:*
em_htg_hum:*
em_htg_inv:*
em_htg_other:*
em_htg_mus:*
em_htg_pln:*
em_htg_rod:*
em_htg_mam:*
em_htg_vrt:*
em_sy:*
em_htgo_hum:*
em_htgo_mus:*
em_htgo_other:*

ALIGNMENTS

RESULT 1
BD062757
LOCUS   BD062757   1168 bp   DNA   linear   PAT 27-
DEFINITION   A tumor necrosis factor related ligand.
ACCESSION   BD062757
VERSION   BD062757.1 GI:22608360
KEYWORDS   JP 2001505407-A/1.
SOURCE   unidentified
  ORGANISM   unidentified
               unclassified.
REFERENCE   1 (bases 1 to 1168)
  AUTHORS   Chicheportiche,Y. and Browning,J.L.
  TITLE   A tumor necrosis factor related ligand
  JOURNAL   Patent: JP 2001505407-A 1 24-APR-2001;
               BIOGEN INC,THE FACULTY OF MEDICINE OF THE UNIVERSITY OF GEN

the number of results predicted by chance to have a

TNF family related protein
JP 2001505407-A/1
24-APR-2001
07-AUG-1997 JP 1998508239
07-AUG-1996 US   60/023541,18-OCT-1996 US   60/028515 PR
R-1997 US   60/040820
VES CHICHEPORTICHE,JEFFREY L BROWNING
C12N15/28,C07K14/525,G01N33/68,C07K16/24,C12N15/11,A61K48/00,
12N5/10,
A61K39/395,A61K38/19,C07K14/705,C12N15/12
Strandedness: Double;
Topology: Linear;
Key   Location/Qualifiers
CDS   2..676.
            /organism="unidentified"
            /mol_type="genomic DNA"
            /db_xref="taxon:32644"

-arity   100.0%;   Score 1168;   DB 6;   Length 1168;
-onservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

| Qy | 721 | CAGGAGCATCACCACACCTCCTACCCCACCCCACTCCTCCACCCCTCGCTGC |
| Db | 721 | CAGGAGCATCACCACACCTCCTACCCCACCCCACTCCTCCACCCCTCGCTGC |

| Qy | 781 | GGTCCAGTCGTGTCTCTCAAAGGCAGCCAGAGCTTGTTCACATGTTTCCATT |
| Db | 781 | GGTCCAGTCGTGTCTCTCAAAGGCAGCCAGAGCTTGTTCACATGTTTCCATT |

| Qy | 841 | GACGTATCCTGCTTCTTAAACATCCACAACTATCCACCTCACTA |
| Db | 841 | GACGTATCCTGCTTCTTAAACATCCACAACTATCCACCTCACTA |

| Qy | 901 | CAAAGCCCCTACTTATCCCTGACTCCCCCACCACTCACCCGACCACGTGTTTAT |
| Db | 901 | CAAAGCCCCTACTTATCCCTGACTCCCCCACCACTCACCCGACCACGTGTTTAT |

| Qy | 961 | TTGTGCACCAGGCACTGAGATGGGCTGAACCTGGTGGCAGGAACCCAGAGAACCT |
| Db | 961 | TTGTGCACCAGGCACTGAGATGGGCTGGACCTGGTGGCAGGAAGCCAGAACCT |

| Qy | 1021 | TAGGCCAGAAGTTCCCAACTGTGAGGGGAAAGAGCTGGGGACAAGCTCCTCCCTG |
| Db | 1021 | TAGGCCAGAAGTTCCCAACTGTGAGGGGAAAGAGCTGGGGACAAGCTCCTCCCTG |

| Qy | 1081 | CTGTGGATTTGAAAAGATACTATTTTATTATTGTGACAAAATGTTAAATG |
| Db | 1081 | CTGTGGATTTGAAAAGATACTATTTTTATTATTGTGACAAAATGTTAAATG |

| Qy | 1141 | TAAAGAGAATAAATCATGATTTCTTC 1168 |
| Db | 1141 | TAAAGAGAATAAATCATGATTTCTTC 1168 |

RESULT 2
AF030100
LOCUS   AF030100   1239 bp   mRNA   linear   ROD 3(
DEFINITION Mus musculus TWEAK mRNA, complete cds.
ACCESSION   AF030100
VERSION   AF030100.2 GI:33348855
KEYWORDS   .
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murir
REFERENCE   1 (bases 1 to 1239)
  AUTHORS   Chicheportiche,Y., Bourdon,P.R., Xu,H., Hsu,Y.M., Scott,H.
            Hession,C., Garcia,I. and Browning,J.L.
  TITLE   TWEAK, a new secreted ligand in the tumor necrosis factor
            that weakly induces apoptosis
  JOURNAL   J. Biol. Chem. 272 (51), 32401-32410 (1997)
  MEDLINE   98070415
  PUBMED   9405449
REFERENCE   2 (bases 1 to 1239)
  AUTHORS   Chicheportiche,Y., Bixler,S., Tizard,R. and Browning,J.
  TITLE   Direct Submission
  JOURNAL   Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge
            Cambridge, MA 02142, USA
REFERENCE   3 (bases 1 to 1239)
  AUTHORS   Chicheportiche,Y., Bixler,S., Tizard,R. and Browning,J.
  TITLE   Direct Submission
  JOURNAL   Submitted (30-JUL-2003) Cell Biology, Biogen, 12 Cambridge
            Cambridge, MA 02142, USA
REMARK   Sequence update by submitter
COMMENT   On Jul 30, 2003 this sequence version replaced gi:2707220.
FEATURES   Location/Qualifiers
     source   1..1239
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /cell_type="peritoneal macrophages"
     CDS   1..750
                     /note="ligand in the TNF family; secreted protein
                     /codon_start=1

| Qy | 60 | CTGAGCCCTGGCCTGGCGCTGCCGCTCCTGGCCTCCTGGTCGTGTCGTCAGCCT |
| Db | 60 | CTGAGCCCTGGCCTGGCGCTGCCGCTCCTGGCCTCCTGGTCGTGTCGTCAGCCT |

| Qy | 120 | AGCTGGGCAACGCTGTCTGCCCCAGGAGCCTTCTCAGGAGGAGCTGACAGCAGAGGA |
| Db | 120 | AGCTGGGCAACGCTGTCTGCCCCAGGAGCCTTCTCAGGAGGAGCTGACAGCAGAGGA |

| Qy | 180 | CGGGAGCCCCTGAACTGAATCCCCAGACAGAGAGAAGCCAGGAGTGTCGGTACCTTT |
| Db | 180 | CGGGAGCCCCTGAACTGAATCCCCAGACAGAGAGAAGCCAGGAGTGTCGGTACCTTT |

| Qy | 240 | GAACAACTAGTCCGGCCTCGAAGAGTGCTCTAAAGGCCGGAAGGGCCGGCCTCG |
| Db | 240 | GAACAACTAGTCCGGCCTCGAAGAGTGCTCTAAAGGCCGGAAGGGCCGGCCTCG |

| Qy | 300 | GCTATTGCAGCCCATTATGAGGTTCATCCTCGGCCAGGACAGGATGGAGCACAAGC |
| Db | 300 | GCTATTGCAGCCCATTATGAGGTTCATCCTCGGCCAGGACAGGATGGAGCACAAGC |

| Qy | 360 | GTGGATGGGACAGTGAGTGGCTGGGAAGAGACCAAAATCAACAGCTCCAGCCCTCT |
| Db | 360 | GTGGATGGGACAGTGAGTGGCTGGGAAGAGACCAAAATCAACAGCTCCAGCCCTCT |

| Qy | 420 | TACGACCGCCAGATTGGGGAATTTACAGTCATCAGTGGGCTTCTACTACCTGTA |
| Db | 420 | TACGACCGCCAGATTGGGGAATTTACAGTCATCAGTGGGCTTCTACTACCTGTA |

| Qy | 480 | CAGGTGCACTTTGATGAGGGAAAGGCTGTCTACCTGAAGCTTGCTGGTGAA |
| Db | 480 | CAGGTGCACTTTGATGAGGGAAAGGCTGTCTACCTGAAGCTTGCTGGTGAA |

| Qy | 540 | GTGCTGGCCCTGCGCGCTGGAAGAATTCTCAGCCACAGCAGCAAGCTCTCCTGG |
| Db | 540 | GTGCTGGCCCTGCGCGCTGGAAGAATTCTCAGCCACAGCAGCAAGCTCTCCTGG |

| Qy | 600 | CAGCTCCGTTGTGCCAGGTGTCTGGGCTGTTGCCGCGGCCAGGGTCTTCCCT |
| Db | 600 | CAGCTCCGTTGTGCCAGGTGTCTGGGCTGTTGCCGCGGCCAGGGTCTTCCCT |

| Qy | 660 | ATCCGCACCCTCCCCTGGGCTCATCTTAAGGCTGCCCCCTTCCTAACCTACTTTGG |
| Db | 660 | ATCCGCACCCTCCCCTGGGCTCATCTTAAGGCTGCCCCCTTCCTAACCTACTTTGG |

| Qy | 720 | TTTCAAGTTCACTGAGGGGCCTTGCTCTCCCAGATTCCTTAAACTTTCCCTGGCTC |
| Db | 720 | TTTCAAGTTCACTGAGGGGCCTTGCTCTCCCAGATTCCTTAAACTTTCCCTGGCTC |

/product="TWEAK"
/protein_id="AAC53517.2"
/db_xref="GI:33348856"
/translation="MAARRSQRRRGRRGEPCTALLAPLVLSLGLALACLGLLLVVVSL
GSWATLSAQEPSQEELTAEDRREPELNPQTEESQDVVPFLEQLVRPRRSAPKGRKAR
PRRAIAAHYEVHPRPGDGAQAGVDGTVSGWEETKINSSSPLRYDRQIGEFTVIRAGL
YYLYCQVHFDEGKAVYLKLDLLVNGVLALRCLEEFSATAASSPGPQLRLCQVSGLIPL
RPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH"

```
rity    100.0%;  Score 1168;  DB 10;  Length 1239;
        100.0%;  Pred. No. 1.6e-298;
nservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

TGAGCCTCGGGCCTGGGCCTGCCTCCTGCCTCGGTCGTGTCAGCCT        60
||||||||||||||||||||||||||||||||||||||||||||||
TGAGCCTCGGGCCTGGGCCTGCCTCCTGCCTCGGTCGTGTCAGCCT       131

GCTGGGCAACGCTGTCTGCCCAGGAGAGCCTTCTCAGGAGGAGCTGACAGCAGAGGA    120
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCTGGGCAACGCTGTCTGCCCAGGAGAGCCTTCTCAGGAGGAGCTGACAGCAGAGGA   191

GGGAGCCCCCTGAACTGAATCCCCAGACAGAGGAAAGCCAGGATGTGTACCTTT      180
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGGAGCCCCCTGAACTGAATCCCCAGACAGAGGAAAGCCAGGATGTGTACCTTT     251

AACAACTAGTCCGGCCTCGAAGAGTGCTCCTAAAGGCCGGAAGGCGGCGGCCTCG     240
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
AACAACTAGTCCGGCCTCGAAGAGTGCTCCTAAAGGCCGGAAGGCGGCGGCCTCG    311

CTATTGCAGCCCATTATGAGGTTCATCCTCGGCCAGGACAGGATGGAGCACAAGC     300
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTATTGCAGCCCATTATGAGGTTCATCCTCGGCCAGGACAGGATGGAGCACAAGC    371

TGGATGGGACAGTGAGTGGCTGGGAAGAGACCAAAAATCAACAGCTCCAGCCCTCT    360
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGGATGGGACAGTGAGTGGCTGGGAAGAGACCAAAAATCAACAGCTCCAGCCCTCT   431

ACGACCGCCAGATTGGGGAATTACAGTCATCAGGGCTGGGCTCTACTACCTGTA     420
|||||||||||||||||||||||||||||||||||||||||||||||||||||
ACGACCGCCAGATTGGGGAATTACAGTCATCAGGGCTGGGCTCTACTACCTGTA    491

AGGTGCACTTTGATGAGGGAAAGGCTGTCTACCTGGACTTGCTGGTGAA         480
|||||||||||||||||||||||||||||||||||||||||||||||||
AGGTGCACTTTGATGAGGGAAAGGCTGTCTACCTGGACTTGCTGGTGAA        551

TGTGGGCCTGGCTGCCTGGAAGAATTCTCAGCCACAGCAGCAAGTCTCCTGG       540
|||||||||||||||||||||||||||||||||||||||||||||||||||
TGTGGGCCTGGCTGCCTGGAAGAATTCTCAGCCACAGCAGCAAGTCTCCTGG      611

AGTCCGTTTGTGCCAGGTGTCTGGGCGTGTTGCCGCTGCGGGCCAGGGTCTTCCCT    600
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGTCCGTTTGTGCCAGGTGTCTGGGCGTGTTGCCGCTGCGGGCCAGGGTCTTCCCT   671

TCCGCACCCTCCCTGGGCTCATCTTAAGGCTGCCCCCTTCCTAAACTACTTTGG     660
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCCGCACCCTCCCTGGGCTCATCTTAAGGCTGCCCCCTTCCTAAACTACTTTGG    731

TTCAAGTTCACTGAGGGGGCCTTGCTCTCCCAGATTCCTTAAACTTTCCCTGGCTC   720
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TTCAAGTTCACTGAGGGGGCCTTGCTCTCCCAGATTCCTTAAACTTTCCCTGGCTC   791

GCATCACCACCACCTCCCTACCCCACCCCACTCTCCACCCCCCTCGCTGTCCTTT    780
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCATCACCACCACCTCCCTACCCCACCCCACTCTCCACCCCCCTCGCTGTCCTTT   851

AGTCCTGTCTCTCCTCAAAGGCAGCCACCAGAGCTTGTTCCACATGTTCCATTCCACA    840
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGTCCTGTCTCTCCTCAAAGGCAGCCACCAGAGCTTGTTCCACATGTTCCATTCCACA   911

ATCCTTGCTTCTTAACATCCCATCCCACCACAACTATCCACCTCACCTCAGCTCC    900
||||||||||||||||||||||||||||||||||||||||||||||||||||||
ATCCTTGCTTCTTAACATCCCATCCCACCACAACTATCCCACCTCACCTCAGCTCC   971
```

```
Qy   901  CAAAGCCCCTACTTATCCCTGACTCCCCACCCACTCACCGACCACCACGTGTTATTC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   972  CAAAGCCCCTACTTATCCCTGACTCCCCACCCACTCACCGACCACCACGTGTTATTC

Qy   961  TTGTGCACCAGGCACTGAGATGGCTGGACCTGTGTGGCAGGAGGCCAGAGAACCTGC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1032  TTGTGCACCAGGCACTGAGATGGCTGGACCTGTGTGGCAGGAGGCCAGAGAACCTGC

Qy  1021  TAGGCCAGAAGTTCCCAACTGTGAGGGGGAAGAGTCGGGACAAGCTCCTCCCTGGA
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1092  TAGGCCAGAAGTTCCCAACTGTGAGGGGGAAGAGTCGGGACAAGCTCCTCCCTGGA

Qy  1081  CTGTGGATTTTGAAAAGATACATATTTTATTATTATTGTGACAAAATGTTAAATGGA
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1152  CTGTGGATTTTGAAAAGATACATATTTTATTATTATTGTGACAAAATGTTAAATGGA

Qy  1141  TAAAGAGAATAAATCATGATTTCCTTC 1168
          |||||||||||||||||||||||||||
Db  1212  TAAAGAGAATAAATCATGATTTCCTTC 1239
```

RESULT 3
AC069459/c
LOCUS   AC069459   203083 bp   DNA   linear   HTG 27-
DEFINITION Mus musculus chromosome 11 clone RP23-168P5, WORKING DRAFT
            SEQUENCE, 7 unordered pieces.
ACCESSION   AC069459
VERSION   AC069459.23 GI:14547768
KEYWORDS   HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleo
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murina
REFERENCE   1 (bases 1 to 203083)
  AUTHORS   Metzker,M.L., Lewis,L.R., Hume,J., Edwards,C., Harris,C.,
            Dederich,D., Thomas,S., Okwuonu,G., Carlock,C., Garner,T.,
            Addison,S., Pace,A., Williams,G., Bonnin,D., Brooks,A., Bro
            Buhay,C., Bunac,C., Burkett,C., Chacko,J., Chen,G., Chen,Z.
            Cox,C., Davis,C., Delgado,O., Ding,Y., Dugan-Rocha,S.,
            Fernandez,C., Ferraguto,D., Forcum-Tansey,J., Gill,R.,
            Gorrell,J.H., Gunaratne,P., Haller,G., Hernandez,J., Hogue
            Hosak,H., Hou,X., Huber,J., Jackson,L., Jia,Y., Kelly,J., H
            Kovar,C., Liu,J., Liu,W., Louiseged,H., Lozado,R.J., Martin
            Massey,E., McLeod,M.P., Mei,C., Moore,S., Morgan,M., Morris
            Neal,D., Nelson,A., Nguyen,R., Nguyen,N., Ogunh,M., Parish,
            Perez,L., Reiter,D., Say,J., Shen,H., Vasquez,L., Watlingto
            Williamson,A., Wrensford,G., Zhou,X., Bouck,J., Hodgson,A.,
            Muzny,D.M., Rives,M., Scherer,S., Sodergren,E., Weinstock,C
            Worley,K. and Gibbs,R.
  TITLE   Direct Submission
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 203083)
  AUTHORS   Worley,K.C.
  TITLE   Direct Submission
  JOURNAL   Submitted (31-MAY-2000) Human Genome Sequencing Center, Dep
            of Molecular and Human Genetics, Baylor College of Medicine
            Baylor Plaza, Houston, TX 77030, USA
COMMENT   On Jun 25, 2001 this sequence version replaced gi:12621364
            Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu/
            Contact: hgsc-help@bcm.tmc.edu
            ---------------- Project Information ----------------
            Center project name: MAFO
            Center clone name: RP23-168P5
            ------------------ Summary Statistics ------------
            Sequencing vector: M13; L08821
            Chemistry: Dye-primer Bodipy: 48% of reads
            Chemistry: Dye-terminator Big Dye: 52% of reads
            Assembly program: Phrap; version 0.990329
            Consensus quality: 212648 bases at least Q40

nsensus quality: 218902 bases at least Q30
nsensus quality: 222384 bases at least Q20
timated insert size: 210656; sum-of-contigs estimation
ality coverage: 0x in Q20 bases; agarose-fp estimation
ality coverage: 7.2x in Q20 bases; sum-of-contigs estimation

: Estimated insert size may differ from sequence length
ee http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
: This is a 'working draft' sequence. It currently
ists of 7 contigs. The true order of the pieces
ot known and their order in this sequence record is
trary. Gaps between the contigs are represented as
of N, but the exact sizes of the gaps are unknown.
record will be updated with the finished sequence
son as it is available and the accession number will
reserved.

```
1       62152:  contig of 62152 bp in length
2153    62252:  gap of unknown length
2253   118872:  contig of 56520 bp in length
8773   118872:  gap of unknown length
8873   148924:  contig of 30052 bp in length
8925   148924:  gap of unknown length
9025   167231:  contig of 18207 bp in length
7232   167331:  gap of unknown length
7332   189907:  contig of 22576 bp in length
9908   190007:  gap of unknown length
0008   196537:  contig of 6530 bp in length
6538   196637:  gap of unknown length
6638   203083:  contig of 6446 bp in length.
```

Location/Qualifiers
1..203083
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-168P5"

arity   60.9%;   Score 711; DB 2; Length 203083;
99.6%;   Pred. No. 5.5e-177;
onservative   0; Mismatches   0; Indels   0; Gaps   3;

```
Qy  TGCACTTTGATGAGGAAAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGAACGGT  484
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  TGCACTTTGATGAGGAAAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGAACGGT  42960

Qy  TGGCCCTGCGCTGCCTGGAAGAATTCTTCAGCCACAGCAGCAAGCTCCTCCTGGGCCC  544
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  TGGCCCTGCGCTGCCTGGAAGAATTCTTCAGCCACAGCAGCAAGCTCCTCCTGGGCCC  42900

Qy  TCCGTTTGTGCCAGGTGTCTGGGCTGTTGCCGCTGCGGCCAGGGTCTTCCCTTCGG  604
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  TCCGTTTGTGCCAGGTGTCTGGGCTGTTGCCGCTGCGGCCAGGGTCTTCCCTTCGG  42840

Qy  GCACCCTCCCCTGGGCTCATCTTAAGGCTGCCCCCTTCCTAACCTACTTTGACTC  664
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  GCACCCTCCCCTGGGCTCATCTTAAGGCTGCCCCCTTCCTAACCTACTTTGACTC  42780

Qy  AAGTTCACTGAGGGGCCTGCTCTCCCAGATTCCTTAAACTTTCCCTGGCCTCCAGG  724
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  AAGTTCACTGAGGGGCCTTGCTCTCCCAGATTCCTTAAACTTTCCCTGGCCTCCAGG  42720

Qy  TCACCACACCTCCCTACCCCACCCCCACCCCCACTCCTCCCACCCCCTCGCTGCCTCCTTCGTC  784
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  TCACCACACCTCCCTACCCCACCCCCACCCCCACTCCTCCCACCCCCTCGCTGCCTCCTTGGTC  42660

Qy  CCTGT-CTCTCCTCAAAGGCCAGAGCTTGTTCACATG-TTTCCATTCCACAGA  842
    |||||  |||||||||||||||||||||||||||||| |||||||||||||||
Db  CCTGTGCTCTCCTCAAAGGCCAGAGCTGTTGTTCACATGTTTCCATTCCACAGA  42600

Qy  TCCTTGCTGCTTC-TTAACATCCATCCACCACAACTATCCACCTCACTAGCTCCC  901
    ||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db  TCCTTGCTGCTTCTTTAACATCCATCCACCACAACTATCCACCTCACTAGCTCCC  42540
```

```
Qy  902    AAAGCCCTACTTATCCCTGACTCCCCACTCACCCGACCACCAGTGTTTATTC
               |||||||||||||||||||||||||||||||||||||||||||||||
Db  42539  AAAGCCCTACTTATCCCTGACTCCCCACTCACCCGACCACCAGTGTTTATTC

Qy  962    TGTGCACCAGGCACTGAGATGGGCTGGACTGGTGGCAGGAAGCCAGAGAACCTGG
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  42479  TGTGCACCAGGCACTGAGATGGGCTGGACTGGTGGCAGGAAGCCAGAGAACCTGC

Qy  1022   AGGCCAGGAAGTTCCCAACTGTGAGGGGGAAGAGTCGGGACAAGCTCCTCCCTGGA
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  42419  AGGCCAGGAAGTTCCCAACTGTGAGGGGGAAGAGTCGGGACAAGCTCCTCCCTGGA

Qy  1082   TGTGGATTTGAAAAGATACATATTTTATTATTATTGTGACAAAATGTTAAATGGA
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  42359  TGTGGATTTGAAAAGATACATATTTTATTATTATTGTGACAAAATGTTAAATGGA

Qy  1142   AAAGAGAATAAATCATGATTTCTCTTC  1168
               ||||||||||||||||||||||||||
Db  42299  AAAGAGAATAAATCATGATTTCTCTTC  42273
```

RESULT 4
AL603707/c
LOCUS   AL603707   234182 bp   DNA   linear   ROD 17
DEFINITION Mouse DNA sequence from clone RP23-422L16 on chromosome 11
            complete sequence.
ACCESSION   AL603707
VERSION   AL603707.5 GI:17017790
KEYWORDS   HTG.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murin
            1 (bases 1 to 234182)
REFERENCE
AUTHORS   Pearce,A.
TITLE   Direct Submission
JOURNAL   Submitted (17-NOV-2001) Wellcome Trust Sanger Institute, F
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.
COMMENT   On Nov 20, 2001 this sequence version replaced gi:1660576...
            During sequence assembly data is compared from overlapping
            Where differences are found these are annotated as variati
            together with a note of the overlapping clone name. Note t
            variation annotation may not be found in the sequence subm
            corresponding to the overlapping clone, as we submit seque
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise not
            regions were either double-stranded or sequenced with an a
            chemistry or covered by high quality data (i.e., phred qua
            30); an attempt was made to resolve all sequencing problem
            as compressions and repeats; all regions were covered by a
            one plasmid subclone or more than one M13 subclone; and th
            assembly was confirmed by restriction digest. The followir
            abbreviations are used to associate primary accession numb
            in the feature table with their source databases: Em:, EME
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the W
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-42
            from the RPCI-23 Mouse PAC Library
            constructed by the group of Pieter de Jong.
            For further details see http://www.chori.org/bacpac/home.h
            VECTOR: pBACe3.6
            This sequence is the entire insert of clone RP23-422L16.
FEATURES   Location/Qualifiers
     source   1..234182
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="11"
                     /clone="RP23-422L16"
                     /clone_lib="RPCI-23"
     misc_feature   complement(84050..84131)

/note="Sequence from uni-directional primer reads and dGTP big dye terminator reads only."

AUTHORS   Ashkenazi,A.J., Goddard,A., Godowski,P.J., Gurney,A.L., Hillan,K.J., Marsters,S.A., Pan,J., Pitti,R.M., Roy,M.A., s Stone,D.M., Watanabe,C.K. and Wood,W.I.
TITLE   Compositions and methods for the treatment of tumour
JOURNAL   Patent: WO 0153486-A 3 26-JUL-2001;
          Genentech, Inc. (US)
FEATURES   Location/Qualifiers
     source   1..1353
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match   53.8%; Score 628.6; DB 6; Length 1353;
Best Local Similarity 76.2%; Pred. No. 2.1e-155;
Matches 933; Conservative 0; Mismatches 219; Indels 73; Ga

rity   60.9%; Score 711; DB 10; Length 234182;
99.6%; Pred. No. 5.7e-177;
nservative   0; Mismatches 0; Indels   3; Gaps   3;

Qy   2 GTGCTGAGCCTGGGCGCTGGCGCCTGGCCTTGGCCTCCTGGTCGTGGTCAGC
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   130 GGGCTGGGCCTGGGCGCTGGGCGCCTGGCCTTCCTGCGTCGGCCGTGGTCAGT

3CACTTTGATGAGGGAAAGGCTGTCTACCTGAACGGCT 484
|||||||||||||||||||||||||||||||||||||
3CACTTTGATGAGGGAAAGGCTGTCTACCTGAACGGCT 66264

Qy   62 GGGAGCTGGGCAACGCTGTCTGCCCAGGAGCCTTCTCAGGAGGAGCTGACAGCAGAG
               ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   190 GGGAGCCGGGCATCGCTGTCGCCCCAGGAGCCTGCCCAGGAGGAGCTGGCAGAG

3GCCCTGCGCTGCCTGGAAGAATTCTCAGCCACAGCAGCAAGCTCTCCTGGGCCC 544
|||||||||||||||||||||||||||||||||||||||||||||||
3GCCCTGCGCTGCCTGCCTGGAAGAATTCTCAGCCACAGCAGCAAGCTCTCCTGGGCCC 66204

Qy   122 CGCCGGAGCCCCCTGAACTGAATCCCCAGACAGAGGAAAGCCAGGATGTGGTACCT
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   250 GACCAGGACCCCGTCGGAACTGAATCCCCAGACAGAGGAAGCCAGGATCCTGCGGCCT

CCGTTTGTGCCAGGTGTCTGGGCTGTTGCCCGCTGCGCCAGGTCTTCCCTTCGG 604
|||||||||||||||||||||||||||||||||||||||||||||||
CCGTTTGTGCCAGGTGTCTGGGCTGTTGCGCCGGTGCGCCAGGGTCTTCCCTTCGG 66144

Qy   182 TTGGAACAACTAGTCCGGCCTCGAAGAGTGCTGGGGGCCTAAAGGCCGGAAGGCGCGCCT
               ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   310 CTTGGAACCGACTAGTTCGGCCTCGCAGAGTGCTCAGCACTCCTCGGGGCGAAAACACGGGCT

CACCCTCCCCTGGCTCATCTTAAGGCTGCCCCCTTCTAACTACTTTGGACTC 664
|||||||||||||||||||||||||||||||||||||
CACCCTCCCCTGGCTCATCTTAAGGCTGCCCCCTTCTAACCTACTTTGGACTC 66084

Qy   242 CGAGCTATTGCAGCCCATTATGAGGTTCATCCTCGGCCAGGACAGGATGGAGCACAA
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   370 AGAGCGGATCGCAGCCCATTATGAAGTTCATCACGACGTCGGGGAAGAGACCAAATCAACAGTTCCAGCCCT

AGTTCACTGAGGGGCCTTGCTCTCCCAGATTCCTTAAACTTTCCCTGGCTCCAGG 724
|||||||||||||||||||||||||||||||||||||
AGTTCACTGAGGGGCCCTTGCTCTCCCAGATTCCTTAAACTTTCCCTGGCTCCAGG 66024

Qy   302 GGTGTGGATGGACAGTGAGTGGGAAGAGACCAAAATCAACAGTCTGGGCTCTCTACTACCTG
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   430 GGTGTGGACGGACAGTGAGTGCTGGGAGGAAGGCGTCGTTACCTACCTGCTGGGCTCTCTACTACCTG

CACCACACCCTCCCTACCCCACCCCACCCCCACCCCCACCTCCTCCACCCTCCACCCTCGCTGGTC 784
|||||||||||||||||||||||||||||||||||||
CACCACACCCTCCCTACCCCACCCCACCCCCACCCCCACCTCCTCCACCCTCCACCCTCGCTTGTC 65964

Qy   362 CGCTACGACCGCCAGATTGGGAATTACAGTCATCAGGGCTGGCTCTCGAAGCTGGACTTGCTGGTG
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   490 CGCTACAACCGCCAGATCGGGAGTTTATAGTCACCCGGGCTGGCTCTCGAAGCTGGACTTGCTGGTG

CTGT-CTCTCCCAAAGGCAGCCAGACCTTGTTCACATG-TTTCATTCCACAGA 842
|||||||||||||||||||||||||||||||||||||
CTGTCCTCTCCTTCAAAGGCAGCCAGACCTTGTTCACATGTTTTCCATTCCACAGA 65904

Qy   422 TGTCAGGTGCACTTTGATGAGGGAAAGGCTGTCGTTGGGAAGAATTCTCAGCCACAGCAGCAAGCTCTCCT
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   550 TGTCAGGTGCACTTTGATGAGGGGAAGGAGGCTGTCTCACCTGGGAAGAATTCTCAGCCACTTGCTGTCGTC

CCTTGCTCTTC-TTAACATCCCATCCCACCACAACTATCCACCCTCACTAGCTCCC 901
|||||||||||||||||||||||||||||||||||||
CCTTGCTCTTCTTAACATCCCATCCCACCACAACTATCCACCCTCACTAGCTCCC 65844

Qy   482 GGTGTGCTGGCCCGTTGTGCCAGGTGTCTGGGCTGTTGCCGCTGGGGCCAGGGTCTTCCC
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   610 GGTGTGCTGGCCCTGCTGGCCCTGGAGCAGTCGCTGGAGGCTGGGGCCAGGGTCCAGGGTCCTC

CCCTACTTATCCCTGACTCCCCCCCCACTCACCCGACCACGTGTTATTGACTT 961
|||||||||||||||||||||||||||||||||||||
CCCTACTTATCCCTGACTCCCCCCACCCACTCACCCGACCACGTGTTATTGACTT 65784

Qy   542 CCCCAGCTCCCGTTGTGCCAGGTGTCTGGGCTGTTGCCGCTGGGCCCAGGGGTCTTCCC
               ||||||||||||||||||||||||||||||||||||||||||||||
Db   670 CCCCAGCTCCGCGCTCCGCCCATCTCAAGGTGTCTGGGCTGTTGTGGCCTGTGCCCGCGGCCCCAGGGGTCTCC

ACCAGGCACTGAGATGGGCTGGACCTGGTGGTGGCAGGGGGGAAGAGCTCGGAGAACCTGGGACT 1021
|||||||||||||||||||||||||||||||||||||
ACCAGGCACTGAGATGGGCTGGACCTGGTGGTGGCAGGGGGGAAGAGCTCGGAGAACCTGGGACT 65724

Qy   602 CGATCCGCACCCTCCCCTGGGCTCATCTTAAGGCTGCCCCCTTCCTAACCTACTTT
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   730 CGGATCCGCACCCTCCCCCTGGGCCATCTCAAGCTGCCCCCTTCCTCACCTACTTC

AGAAGTTCCCAACTGTGAGGGGGAAGAGCTGGGGAAGAGCTCCTCCCTGGGATCCC 1081
|||||||||||||||||||||||||||||||||||||
AGAAGTTCCCAACTGTGAGGGGAAGAGCTGGGGAAGAGCTCCTCCCTGGGATCCC 65664

Qy   662 CTCTTTCAAGTTCACTGAGGGGCCTTGCTCTCCCGAGATTCCTTAAACTTTCCCTGGC
               ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   790 CTCTTCCAGGTTCACTGAGGTGGTCCCCGTCTCCCGAGTCGTCCCCAGGCTGCCCAGGCTGCCGGC

ATTTTGAAAAGATAACTATTTTTATTATTATTGTGACAAAATGTTAAATGGATATT 1141
|||||||||||||||||||||||||||||||||||||
ATTTTGAAAAGATAACATATTTTTATTATTTTATTATTATTGTGACAAAATGTTAAATGGATATT 65604

Qy   722 AGGAGCATCACCACCACCCTACCCCACCCCACCCCCGGTCCCCCTGCGCCCCCACCCTCG-GCTGCTC
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   850 ------CCTCGACAGCTCTCTGGGCACCCCGGTCCCCTCTGCCCCACCCTCAGCCCGCTC

GAATAAATCATGATTTCTCTTC 1168
||||||||||||||||||||||
GAATAAATCATGATTTCTCTTC 65577

Qy   781 GGTCCAGTCCTGTCTCCC--TCAAAGGCAGCCAGAGCTTGTTGTTCACATGTTTCCATT
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   905 GCTCCAGACCTGCCTCCCTCCCTCCAGAGCTGCTGGGCGTTGCTGCTGTTCACGTGTTTCAGGTGTTTTTCCAT

4   3 from Patent WO0153486.
e   3 piens (human)
4   piens
4.1   GI:15391154

1353 bp   DNA   linear   PAT 30-AUG-2001

ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```
/product="Apo-3 Ligand"
/protein_id="AAQ89229.1"
/db_xref="GI:37182858"
/translation="MAARRSQRRRGRRGBPGTALLVPLAIGLGLALACI
GSRASLSAQPEAQELVAEEDQPSELNPQTEESQDPAPFLNRLVRPRF
ARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYNRQIC
YYLYCQVHFDEGKAVYLKLDLLVDGVLALRCLEEFSATAASSLGPQLRI
RPGSSLRITLPWAHLKAAPFLTYGLFQVH"
```

ORIGIN

```
Query Match        53.8%;  Score 628.6;  DB 9;   Length 1353;
Best Local Similarity   76.2%;  Pred. No. 2.1e-155;
Matches  933;  Conservative   0;  Mismatches  219;   Indels  73;   C
```

```
Qy      2 GTGCTGAGCCTGGCCTGGCGCTGGCCTTCTGCCCTCTGCCTTCGTGTCGGTGTCAG
          |||||||||||||||| || | || ||||||||||||||||||||||||||||
Db    130 GCGCTGGGCCTGGGCCTGGCGTCGGCCTGGCCTCGGCCTTCGTGTCGGCCGTGGTCAG

Qy     62 GGGAGCTGGGCAACGCTGTCTGCCCAGGAGCTTCTCAGGAGGAGCTGACAGCAGAT
          ||||| ||  | | ||||||  |||||||||||||||||||||||||||||||
Db    190 GGGAGCGGGGCATCGCTGTCTGCCCAGGAGCCTGCCCAGGAGGAGCTGTGGCAGAT

Qy    122 CGCCGGGAGCCCCCTGAACTGAATCCCAGACAGAGAGGAAAGCCAGGATGTGTACC
          ||||| || |||||||||||||||||||||||||||||||||||||||||||
Db    250 GACCAGGACCCGTCGGAACTGAATCCCAGACAGAGAAGCCAGGAATCCTGCGCGCC

Qy    182 TTGGAACAACTAGTCCGGCCTGAAGAAGTGTTCCTAAAGGCCGGAAGGCGCGGCCC
          | |||||||| || |||||||||||||||||||||||||||||||||||||||||
Db    310 CTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAACACGGGC

Qy    242 CGAGCTATTGCAGCCCATTATGAGGTTCATCCTCGGCCAGGACGATGGGAGCAC/
          | | ||||||||| ||||||| || |||||||||||||||||||||||||||
Db    370 AGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTTGGACAGCAGGACGAGCGC/

Qy    302 GGTGTGGATGGGACAGTGAGTGGCTGGGGAAGAGACCAAAAATCAACAGCTCCAGCC/
          |||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db    430 GGTGTGGACGGGACAGTGAGTGGCTGGGGAAGTGCAGAATCAACAGCTCCAGCC/

Qy    362 CGCTACGACCGCCAGATTGGGAATTTACAGTCATCAGGCTGGCTCTACTACC/
          ||||| ||||||||||||||| ||| ||| ||||||| |||||||||||||||
Db    490 CGCTACACCGCCAGATCGGGAGTTTATAGTCACCCGGGCTGGCTCTACTACC/

Qy    422 TGTCAGGTGCACTTTGATGAGGGAAAGGCTGTCACTGACCTGGAGTGGACTTGCTGG/
          || ||||||||||| |||||||||||||| || |||| ||||| |||||||| |||||
Db    550 TGTCAGGTGCACTTTGATGAGGGGAAGGCTGTCACCTGAAGCTGAAGCTGGACTTGCTGG/

Qy    482 GGTGTGCTGGCCCTGCCCTGCCTGGAGAATTCTCAGCCACAGCAAGCTCTC/
          |||||||| ||| || ||| ||||||||| ||||| |||||||| |||||||||
Db    610 GGTGTGCTGGCCCTGCCCTGCCTGGAGAATTCTCAGCCACTCGCGGGCGAGTTCCC/

Qy    542 CCCCAGCTCCGTTTGTGCCAGGTGTCTCGGCGTGTTGCCCGTGCCGCCAGGGTCTT/
          |||||| |||||| |||||| || |||||| ||||| |||||| |||||||||||||
Db    670 CCCCAGCTCCGCGCCTGCCAGGTGTCTCGGCGTGTTGCCCGTGCCCGGGGCCAGGGTCCT/

Qy    602 CGGATCCGCACCCCCCCCGGGCGGGGGGCCTTGGGCTCATCTTAAGGCTGCCCCTTCCTAACCTACT/
          |||||||||||| || ||||||||||| ||||||||||||||||||||||||||||||||||||
Db    730 CGGATCCGCACCCCCCCCGGGCCCATCTCAAGGCTGCCCATCTCAAGGCTGCCCCTTCCTCACTACT/

Qy    662 CTCTTCAAGTTCACTGAGGGGGCCTTGCTTCTCCCAGATTCCTTAAACTTTCCCTG/
          ||||||||||| || |||| ||||| ||||||||||||||||| || |||||| |||||
Db    790 CTCTTCAGGTTCACTGAGGGGGCCTTGCTCTCCCCGCAGTCGTCCCAGCGTGCGCG/

Qy    722 AGGAGCATCACCACACTCCCTACCCCACCCCCACCCCCTC-GCTGC/
          |||||||||| |||||||||||||||||||  ||||||||  |||||
Db    850 ----CCTCGACAGCTCCTGGGACCCGGTCCCCTGCCCCCACCCCTCAGCCGC/

Qy    781 GGTCCAGTCCTGTCTCTCC--TCAAAGGCAGCCAGAGCTTGTTCACATGTTTCCA/
          ||||||||||||||  |||||||||| ||| |||||||||||||||||||
Db    905 GCTCCAGACCTGCCCCTCCCTCTAGAGGTGCCTGCTGGGCCTGTTCACGTGTTTCCJ

Qy    838 ------ACAGACGTATCCTTGCTCCTTCTTAACATCCCACCACAACTATCCJ
                 ||||||||| || || || |||| |||||||||||||||||||||||||
Db    965 ACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCACTCTCCJ
```

Qy --ACAGACGTATCCTGCTCTCTTAACATCCCATCCACCACAACTATCCACTC 891
    |||||| || || |||||||||||| || ||||||||||||||
Db AAATACAGTATTCCACTCTTATCTTACAACTCCCCACCGCCACTCTCCCACCTC 1024

Qy GCTCCCAAAGCCCCTAC---------TTATCCTGACTCCCCCCACCCACT 936
    |||||||| |||||| |||||||||||||||
Db GCTCCCCAATCCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCCCCCCTGGCCA 1084

Qy CGACCACGTGTTTATTGACTTGTGCAC------------------------- 968
    ||||||||||||||||||| ||||||
Db CCCCAGGTCATTGTGTTCACTGTACTCTGTGTGGGCAAGGATGGTCCAGAAGACCCC 1144

Qy -CAGGCACTGAGATGGGCTGGACCTGGTGGCGGGCGCAGGGGGCGGAGCCCAGAGAACCTGGGACTAG 1023
   |||||||||||||||||||||| |||| ||||||||||||||||||||||||||||||||
Db TCAGGCACTAAGAGGGGCTGGACCTGGTGGCGCGGCGCAGGGGGCGGAGGCCCAGAGAACCTGGGCCTAG 1204

Qy GAAGTTCCCAACTGTGAGGGGGAAGAGCTGGGGACAAGCTCCTCCCTGA----TC 1079
    |||||||||||||||||||||||||||||||||||||||||||||||
Db GGAGTTCCCAAATGTGAGGGGGCGAGA-AACAAGACAAGCTCCTCCCTTGAGAATTC 1263

Qy TTGGATTTTGAAA--AGATACTATTTTTATTATTATTGTGACAAAATGT---TAAAT 1134
    |||||||||||||| ||||||| ||||||||||||||||||||||||
Db TTGGATTTTAAAACAGATATTTTTATTTATTATTATTGTGACAAAAATGTTGATAAAT 1323

Qy AATTAAAGAGAATAAATCATGA 1159
    |||||||||||||||||
Db AATTAAAATGAATAAGTCATAA 1348

`70   1353 bp   mRNA   linear   PRI 03-OCT-2003
apiens clone DNA30879 Apo-3 Ligand (UNQ181) mRNA, complete

`70
`70.1   GI:37182857
NA.
apiens
apiens (human)
oa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 1353)
H.F., Gurney,A.L., Abaya,E., Baker,K., Baldwin,D., Brush,J.,
., Chow,B., Chui,C., Crowley,C., Currell,B., Deuel,B.,
., Eaton,D., Foster,J., Grimaldi,C., Gu,Q., Hass,P.E.,
s., Huang,A., Kim,H.S., Klimowski,L., Jin,Y., Johnson,S.,
Lewis,L., Liao,D., Mark,M., Robbie,E., Sanchez,C.,
feld,J., Seshagiri,S., Simmons,L., Singh,J., Smith,V.,
n,J., Vagts,A., Vandlen,R., Watanabe,C., Wieand,D., Woods,K.,
d,A., Yansura,D., Yi,S., Yu,G., Yuan,J., Zhang,M., Zhang,Z.,
d,A., Wood,W.I. and Godowski,P.
screted Protein Discovery Initiative (SPDI), a Large-Scale
to Identify Novel Human Secreted and Transmembrane Proteins:
nformatics Assessment
Res. 13 (10), 2265-2270 (2003)
`09
H.F.
Submission
ted (01-AUG-2003) Department of Bioinformatics, Genentech,
1 DNA Way, South San Francisco, CA 94080, USA
Location/Qualifiers
    1..1353
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="DNA30879"
    1..1353
    /locus_tag="UNQ181"
    58..807
    /locus_tag="UNQ181"
    /note="PRO207"
    /codon_start=1
```

RPGSLRIRTLPWAHLKAAPFLTYRGLFQVH"

ORIGIN

Query Match          53.8%;    Score 628.6;  DB 9;   Length 1368;
Best Local Similarity 76.2%;   Pred. No. 2.1e-155;
Matches 933;   Conservative 0;  Mismatches 219;  Indels 73;   Ga

Qy     2  GTGCTGAGCCTGGGCCTGGGCCTGGGCGCTGGCCTTCTCCTGCTGGTCTGGTGGTCAGC
Db   130  GCGCTGGGCCTGGGCCTGGGCCTGGGCGCTGGCCTTCTCCTGCTGGCCGTGGTGCAGT

Qy    62  GGGGAGCTGGGCAACGCTGTCTCGCCAGGAGCCTTCTCAGGAGGAGCTGACAGCAGAG
Db   190  GGGGAGCCGGGCATCGCGTCCGCCCAGGAGCCCAGGAGCCTGGTGCGTGTGGCCAGAG

Qy   122  CGCCGGGAGCCCCGTCGAACTGAATCCCAGACAGAGAAGAGAAGCCAGGATGTGGTACCT
Db   250  GACCAGGACCCGTCGGAACTGAATCCCAGACAGAGAAGAAAGCCAGGATCCTGCGCCT

Qy   182  TTGGACAACTAGTCCGGCCTCGAAGAAGTGCTTCCTAAAGGCCGGAAGGCGCGGCCT
Db   310  CTGAACCGACTAGTTCCGCCTCGCAGAAGTCATCCCTCGGCCAGGACAGGAGGATGGAGCACAA

Qy   242  CGAGCTATTGCAGCCCATTAGAGTTCATCCTGGGAAGAGACCAAAATCAACAGCTCCAGCCCT
Db   370  AGAGCGGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACACGACCAGGAGCGGCAG

Qy   302  GGTGTGATGGGACAGTGAGTGGCTGGGAAAGACCAAAATCAACAGCTCCAGCCCT
Db   430  GGTGTGAGCGGGACAGTGAGTGCTGGGAGGAAGCCAGAATCAACAGCTCCAGCCCT

Qy   362  CGCTACGACCGGCCAGATTGGGGAATTTACAGTCATCAGGGCTGGCTTCACTACCTG
Db   490  CGCTACAACCGGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGCTTCACTACCTG

Qy   422  TGTCAGGTGCACTTGATGAGGGAAAGCCTGTCTACCTGAAGCTGGACTTGCTGGTG
Db   550  TGTCAGGTGCACTTTGATGAGGGAAGCCTGTCTACCTGAAGCTGGACTTGCTGGTG

Qy   482  GGTGTGCTGGCCCTGCCCTGCCGCCTGCCGGAAGAATTCTCAGCCACAGCAGCAAGCTCCCT
Db   610  GGTGTGCTGGGCCCTGCCCTGCCGCCTGCCGGAGGAATTCATCAGCCACTGCGGCGGAGTTCCCTC

Qy   542  CCCCAGCTCCGTTGTGCCCAGGTGTCTCGGGCTGTTGCGCTGGGCTGGGCCAGGGTCTTCC
Db   670  CCCCAGCTCCGCCTCGCCTGCCGCCAGGTGTCTCGGGCTGTTGCGCCCTGCCGGCCAGGGTCCTCC

Qy   602  CGGATCCGCACCCTCCCCTGGGCTCATCTTAAGGCTGCCCCCTTCCTAACCTACTTT
Db   730  CGGATCCGCACCCTCCCCTGGGCCATCTCAAGGCTGCCCCCTTCCTCCCACCTACTTC

Qy   662  CTCTTTCAAGTTCACTGAGGGGGGCCTGCTCTCCCAGATTCCTTAAACTTTCCCTGGC
Db   790  CTCTTTCCAGGTTCACTGAGGGGGGCCTGGTCTCCCGCAGTCGTCCCAGGCTGCCGGG

Qy   722  AGGAGCATCACCACAACTCCTACCCCCACTCCTCCACCCCTC-GCTGCTC
Db   850  -----CCTCGACAGCTCTCTGGGCACCCGGTCCCTCTGCCCACCCTCTGCCCACCCTCTCAGCCGCTC

Qy   781  GGTCCAGTGCTCGTCTCC--TCAAAGGCAGCCAGAGCTTGTTCACATGTTTCCATT
Db   905  GCTCCAGACCTGCCCTCCCCTGGGCCCCCTAGAGAGCTGCCTGGGCTGTTCACGTGTTTTCCAT

Qy   838  ------ACAGACGTATCCTTGCTCTTCTTAACATCCCATCCCCACCACACTATCCAC
Db   965  ACATAAATACAGTATCCCACTCTTATCTTACAACTCCCCACTCTCCAC

Qy   892  ACTAGCTCCCAAAGCCCCTAC-----------TTATCCCTGACTCCCCCACCC
Db  1025  ACTAGCTCCCCAATCCTGACCCTTTGAGGGCCCCAGTGATCTCGACTCCCCCCTGC

Qy   937  CACCCGACCACGTGTTTATTGACTTTGTGCAC
           ||||  |||| |||||||

(partial left column, rotated:)

TCCCAAAGCCCCTAC------------TTATCCCTGACTCCCCCACCCCACT 936
TCCCCAATCCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCCCCCTGGCCA 1084

GACCACGTGTTTATTGACTTTGTGCAC----------- 968

CCCCAGGTCATTGTGTTCACTGTACTCTGTGGGCAGGGATGGGTCCAGAGACCC 1144

CAGGCACTGAGGATGGCTGACCTGGTGGCAGGGAGCCAGAGAACCTGGGACTAG 1023
CAGGCACTCAAGAGGGGCTGACCTGGGCCGGCAGGAGCCCAAAGAGACTGGGCCTAG 1204

AAGTTCCAACTGTGAGGGGGGAAGAGTCGGGGACTGGGGACACAAGCTCCTCCCTGGA----TC 1079
GAGTTCCCAAATGTGAGGGGCGGAGA-AACAAGACAAGCTCCTCCCTTGAGAATTC 1263

GGATTTTGAAA--AGATACTATTTTTATTATTATTGTGACAAAATGT---TAAAT 1134
GGATTTTAAAACAGATATTATTTTTATTATTATTGTGACAAAATGTTGATAAAT 1323

TTAAAGAGAATAAAATCATGA 1159
TTAAAATAGAATAAGTCATAA 1348

2        1368 bp  mRNA          linear    PRI 04-MAY-1998
piens Apo3/DR3 ligand (APO3L) mRNA, complete cds.
2
2.1   GI:3108230

piens (human)
ca; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 1368)
s,S.A., Sheridan,J.P., Pitti,R.M., Brush,J., Goddard,A. and
zi,A.
ication of a ligand for the death-domain-containing receptor

iol. 8 (9), 525-528 (1998)
5

es 1 to 1368)
s,S.A., Sheridan,J.P., Pitti,R.M., Brush,J., Goddard,A. and
zi,A.
Submission
ed (25-MAR-1998) Molecular Oncology, Genentech, 1 DNA Way,
an Francisco, CA 94080, USA
Location/Qualifiers
1..1368
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17p13"
1..1368
/gene="APO3L"
58..807
/gene="APO3L"
/function="binds to the death-domain-containing receptor
Apo3/DR3"
/codon_start=1
/note="TNF homolog"
/product="Apo3/DR3 ligand"
/protein_id="AAC39724.1"
/db_xref="GI:3108231"
/translation="MAARRSQRRRGRRGEPGTALLVPLAIGLGLALACLGLLLAVVSL
GSRASLSAQEPAQEELVAEEDQPSELNPQTEESQDPAPPLNRLVRPRRSAPKGRKTR
ARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGL
YYLYCQVHFDEGKAVYLKLDLLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLAL

```
Qy   242  CGAGCTATTGCAGCCCATTATGAGTTCATCTCGGCCAGGACAGGATGAGCACA
               ||||||||||||||||||||||| |||||||||||||||| |||||||||||
Db   404  AGAGGGATCGCAGCCCATTATGAAGTTCATCCACGACGACCTGGACAGGACGAGGAGGCGCA

Qy   302  GGTGTGATGGGACAGTGAGTGGCTGGCTGGGAAGAGACCAAAATCAACAGCTCAGCCC
               |||||||||||||||||||||| |||||||||||||||| |||||||||||||||
Db   464  GGTGTGACGGGACAGTGAGTGGCTGGGACGGGAAGCCAGAATCAACAGCTCCAGCCC

Qy   362  CGCTACGACCGCCAGATTGGGGAATTACAGTCATCAGGGCTGGGCTTCTACTACCT
               ||||||||||||||||||| ||||||||||||||||||||||||| |||||||||
Db   524  CGCTACAACCGCCAGATCGGGAGTTTATAGTCACCCGGGCTGGGCTTCTACTACCT

Qy   422  TGTCAGGTGCACTTTGATGAGGGAAAGGCTGTCTACCTGAAGCTGACTTGCTGGT
               |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   584  TGTCAGGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGACTTGCTGGT

Qy   482  GGTGTCTGGGCCTGCCTGGCTGTCAGGTGTCTGGGCGTGTTGGCCTGCGCTGCTTC
               |||||||||||||||||||||||||||||||||||||||||||||||||
Db   644  GGTGTCTGGGCCTGCCTGGCGTGCTGCCTGGAGCAATTCTCAGCCACTGCGGCGAGTTCCCT

Qy   542  CCCCAGCTCCGTTGTTGCCCAGGTGTCTGGGCGTGTTGGCCGCTGCCAGGGTCTTC
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   704  CCCCAGCTCCGCCCTGCCCGTGTCTGGGCGTGTTGGGCCTGGCCTGCCCTGCGGCCAGGGTCCTC

Qy   602  CGGATCCGCACCCTCCCCTGGGCTCATCTAAGGCTGGCCCCCTTCCTAACCTACTT
               |||||||||||||||||||||||||||||||||||||||||||||||||
Db   764  CGGATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACCTACTT

Qy   662  CTCTTTCAAGTTCACTGAGGGCCTTGCTCCTCCCGAGATTCCTTAAACTTTCCTGG
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   824  CTCTTCCAGTTCACTGAGGGCCCTGCCTCTCCCCGCAGTGTCCCCAGGCTGCCGGC

Qy   722  AGGAGCATCACCACACCTCCCTACCCCACCCCACCTCCTCCACCCCTC-GCTGCT
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   884  -----CCTCGACAGCTCTCGGGCACCCGGTCCCCTCGCCCCTCCTCCAGCCGCT

Qy   781  GGTTCCAGTTCCTGTCTCC--TCAAAGGCAGCCAGAGCTTGTTCACATGTTTCCAT
               |||||||||||||       |||||||||||||||||||||||||||||
Db   939  GCTCCAGACCTGCCCCTCCCCTCTAGAGCGTGCGGCGGCCTGTTCACGTGTTTTCCA

Qy   838  ------ACAGACGTATCCTTGCTCTTCTTAACATCCCATCCACCACCAACATACCA
                     ||||||||||||||||||||||||||| || |||||||||||||
Db   999  ACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCCACCGCCCACTCTCCA

Qy   892  ACTAGCTCCCAAAGCCCCTAC------------TTATCCCTGACTCCCCCACC
               |||||||||||||||              |||||||||||||||||
Db   1059 ACTAGCTCCCAATCCCCTGACCCTTGAGGCCCCCAGTGATCTCGACTCCCCCCTG

Qy   937  CACCCGACCACGTGTTTATTGACTTTGTGCAC-------------------
               |||||||||||||||||||||||| || |||
Db   1119 CAGACCCCCAGGTCATTGTGTCACTGTACTTCTGTGGGCAAGGATGGGTCCGAGAAG

Qy   969  --------CAGGCACTGAGATGGCGGACCTGGACCTGGTGGGAGGAAGCAGAGAACCTGGGA
                        |||||||||||||||||||| ||||||||||||| || |||
Db   1179 CACTTCAGGCACTCAGGGCTGGGACCTGGACCTGGTGGGACAGAGCTGGGA

Qy   1024 GCCGAGAGTTCCAACTGTGAGGGGGAAGAGCTGGGACACCCTG--
               ||||||||||| |||  | || |||||||||||||||||||
Db   1239 GCCGAGGAGTTCCCAAATGTGAGGGGGCGAGA-AACAAGACAAGCTCCTCCCTTGAGA

Qy   1080 CCTGTGGATTTTGAAA--AGATACTATTTTTATTATTATTGTGACAAAATGT----T
               ||||||||||||  || | || ||| ||||||||| || | ||
Db   1298 CCTGTGGATTTTTAAAACAGATATTATTTTTATTATTTTATTGTGACAAAATGTTGAT

Qy   1135 GGATATTAAAGAGAATAAATCATGA 1159
               ||||||||||||||||||||||||
Db   1358 GGATATTAAATAGAATAAGTCATAA 1382
```

```
RESULT 9
AF030099
LOCUS        AF030099              1306 bp    mRNA      linear    PRI 20
DEFINITION   Homo sapiens TWEAK mRNA, complete cds.
```

```
Db   CCCCCAGGTCATTGTGTTCACTGACTCTGTGGGCAAGGATGGGTCCAGAAGACCC 1144
          |||||||||||||||||||||||||||||||||||||||||||||
Qy   -CAGGCACTGAGATGGCGGACCTGGTGGCGAGCAGGAACCTGGGACTAG 1023

Db   TCAGGCACTAAGGAGGGGGCTGGACCTGGCGCGGCAGGAGAAGCCAAAGAGACTGGGCCTAG 1204
          |||||||||||||||||||||

Qy   GAAGTTCCCAACTGTGAGGGGGAAGAGCTGGGACAAGCTCCTCCCGGA----TC 1079
          ||||||||||||||||||||||||||||||||||||||||||
Db   GGAGTTCCAAATGTGAGGGGGCGAGA-AACAAGACAAGCTCCTCCCTGAGAATTC 1263

Qy   TGGATTTTGAAA--AGATACTATTTTTATTATTATTGTGACAAAATGT---TAAAT 1134
          ||||||||||||  ||||||||||||||||||||||||
Db   TGGATTTTTAAAACAGATATTTTATTATTTTATTTGTGACAAAATGTTGATAAAT 1323

Qy   ATTAAAGAGAATAAATCATGA 1159
          |||||||||||||||||
Db   ATTAAATAGAATAAGTCATAA 1348
```

```
52
ligand polypeptide.                   1421 bp    DNA      linear    PAT 27-AUG-2002
52
52.1   GI:22636562
1522584-A/1.
apiens (human)
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 1421)
azi,A.J., Marsters,S.A. and Pitti,R.
ligand polypeptide
; JP 2001522584-A 1 20-NOV-2001;
ECH INC
omo sapiens (human)
P 200152584-A/1
0-NOV-2001
9-OCT-1998 JP 2000516042
0-OCT-1997 US   60/062037,17-DEC-1997 US        60/069862 PI
ASHKENAZI,SCOT A MARSTERS,ROBERT PITTI
12N15/09,A61K38/00,C07K14/705,C07K16/24,C12N15/00,A61K37/02
po-3 ligand polypeptide
ey       Location/Qualifiers
          1..1421
         /organism='Homo sapiens (human)'.
ource    1..1421
         /organism="Homo sapiens (human)"
Location/Qualifiers
1..1421
         /organism="Homo sapiens"
         /mol_type="genomic DNA"
         /db_xref="taxon:9606"
```

```
arity      53.8%;   Score 628.6;   DB 6;   Length 1421;
           76.2%;   Pred. No. 2.1e-155;
onservative  0;   Mismatches 219;   Indels 73;   Gaps 10;
```

```
TGAGCCTGGGCCTGGCGCTGGCCTGCTCCTGCTGCTGGTCGTGTGGTCAGCCTG 61
      ||||||||||||||||||||||||||||||||||||||||||
TGGGCCTGGCGCCGCGCTGGCCTGCTCCTGCTGCTGGTCGTGTGGTCAGTTTG 223

GCTGGGCAACGCGTGTCTGCCCAGGAGCCTTCTCAGGAGGAGTCGACACGCAGAGGAC 121
      |||||||||||||||||||||||||||||||||||||||||
GCCGGGCATCGCGTGTCCGCCCAGGAGCCTGCCCCAGGAGGAGTCGTGTGGCGAGGAG 283

GGGAGCCCCCTGAACTGGCCGAATCCCCAGACAGAGAAAGCCAGGATGTGTACCTTTC 181
      |||||||||||||||||||||||||||||||||| ||||||||||
AGGAACCCCGTCGGAACTGGCCGAATCCCCAGACAGAGAAAGCCAGGATCCTGCGCCTTTC 343

AACAACTAGTCCGGCCTCGAAGAAGTGCTCCTAAAGGCCGAAGGCGCGGCCTCGC 241
      ||||||||||||||||||||||||| || || |||||||||||||||||||
ACCGACTAGTTCCGGCCTCGCAGAGTGCACCTAAAGGCCGAAAACACGGGCCTCGA 403
```

9.
9.1   GI:2707218

piens (human)
piens
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 1306)
ortiche.Y., Bourdon,P.R., Xu,H., Hsu,Y.M., Scott,H.,
,C., Garcia,I. and Browning,J.L.
a new secreted ligand in the tumor necrosis factor family
akly induces apoptosis
. Chem. 272 (51), 32401-32410 (1997)
5

es 1 to 1306)
.P., Hession,C., Tizard,R. and Browning,J.
Submission
ed (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center,
ge, MA 02142, USA
Location/Qualifiers
1..1306
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17p13"
/tissue_type="tonsil"
18..767
/note="ligand in the TNF family; secreted protein; start
codon not verified experimentally"
/codon_start=1
/product="TWEAK"
/protein_id="AAC51923.1"
/db_xref="GI:2707219"
/translation="MAARRSQRRRGRRGEPGTALLVPLALGLGLALACLGLLLAVVSL
GSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTR
ARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGL
YYLYCQVHFDEGKAVYLKLDLLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLAL
RPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH"

rity   53.4%;   Score 624; DB 9; Length 1306;
76.0%;   Pred. No. 3.4e-154;
nservative 0; Mismatches 220; Indels 73; Gaps 10;

GAGCCTGGGCGCTGGGCCTGGCGCTGGCGCTGCTCCTGCGCTGGTCGGTCAGCCTG   61
||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
ACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAACACGGGCTCGA   329

CTATTGCAGCCCATTATGAGGTTCATCCTCCGCCAGGACGATGAGCACAAGCA   301
|||||||||||||||||||||||| |||||||||||||| ||||| |||||
CGATCCGCAGCCCATTATGAGGTTCATCCACGACCTGACAGGAGGACGCCAGGCA   389

TGGATGGACAGTGAGTGGCTGGGAAGAGACCAAAATCAACAGCTTCAGCCCTCTG   361
|||||||||||||||| ||||||||||| ||||| ||||||||||||||||||
TGGACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTTCAGCCCTCTG   449

ACGACCGCCAGATTGGGGAATTTACAGTCATCAGGGCTGGGCTTCTACTACCTGTAC   421
|||||||||||||||||||||| |||||| ||||||||||||||||||||||||
ACAACCCGCCAGATCGGGGAGTTTATAGTGCACCCGGGCTGGGCTTCTACTACCTGTAC   509

Qy 422 TGTCAGGTGCACTTTGATGAGGAAAGGCTGTCTACCTGGAGACTGGACTTGCTGGT(
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 510 TGTCAGGTGCACTTTGATGAGGAAAGGCTGTCTACCTGGAGACTGGACTTGCTGGT(

Qy 482 GGTGTGCTGCGCCCTGGCCTGGAAGAATTCTCAGCCACCAGCAGCAAGCTCTCC.
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db 570 GGTGTGCTGCGCCCTGGCCTGGAAGAATTCTCAGCCACCAGCAGCAAGCTCTCC(

Qy 542 CCCCAGCTCCGTTGTGCCAGTGTCTGGCTGTTGCCGCTGCGGCCCGGCCAGGGTCTTC(
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 630 CCCCAGCTCCGCCTCGCCAGGTGTCTGGGCTGTTGGCCCTGCGGCCCGGCCAGGGTCTC(

Qy 602 CGGATCCGACACCCTCCCCTGGGCTCATCTTAAGGCTGCCCCTTCCTAACCTACT
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 690 CGGATCCGACACCCTCCCCTGGGCCATCTCAAGGGCTGCCCCCTTCCTCACCTACT

Qy 662 CTCTTTCAAGTTCACTGAGGGCCCTTGCTCTCCCAGATTCCTTAAACTTTCCCTGG
              ||||||||||||||||||||||| |||||||||||||| | ||||||||
Db 750 CTCTTCCAGGTTCACTGAGGGCCCTTGCTCTCCCCACAGTCGTCCAGGCTGCCGG(

Qy 722 AGGAGCATCACCACCACCTCCTACCCCACCCCACTCCTCCAC CCCCCTC-GCTGCTI
              ||||||||||||||||||||||||||||||||| |||| ||||||||||
Db 810 ----CCTGACAGTCTCTGGGCACCCGTTCCCCTGCCCCACCCTCAGCCGCGT(

Qy 781 GGTCCAGTCCTGTCTCCC--TCAAAGGCAGCCAGAGCTTGTTGTTCACATGTTCCAT
              |||||||||| || || | |||||||||||||||||||||||||||||||
Db 865 GCTCCAGACCTGCCCCTCCCTCTCTAGAGGGTGCTCTGGGCCTGTTCACGTGTTTTCCA'

Qy 838 ------ACAGACGTATCCTTGCTCTTCTTAACATCCCATCCCACCAACTATCCA
                        ||||||||||||||||||||| |||||||||||||||||
Db 925 ACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCCACTCTCCA

Qy 892 ACTAGCTCCCAAAGCCCCTAC-----------TTATCCCTGACTCCCCCACC
              ||||||||||||||||||||||| ||||||||||||||||||
Db 985 ACTAGCTCCCCAATCCCTGACCCCTTTGAGGCCCCAGTGATCTCGACTCCCCCCTG

Qy 937 CACCCGACCACGTGTTTATTGACTTTGTGCAC-------------------- -T
              |||||||||||||| ||||||||||||||||
Db 1045 CAGACCCCCAGGGCATTGTTCACTGTACTCTGTGGGCAAGGATGGGTCAGAAG

Qy 969 ------CAGGCACTGAGATGGGCTGGACCTGGTGGGCAGGAGAGCCAGAGAACCTGGGA
                       |||||||||||||||||||||||||||| |||||||||||||||||
Db 1105 CACTTCCAGGCACTGACTAAGAGAGGGGCTGGACCTGGGACCTGGGCAGGAGGAGCCAAAGAGACTGGGC

Qy 1024 GCCAGAAGTCCCAACTGTGAGGGGGAAAGCTGGGGACAAGCTCCTCCCTGGA--
              |||||||||||||||||||||| |||||| |||||||||| |||||||||
Db 1165 GCCAGGAGTTCCCAAATGTGAGGGGGCGAGA-AACAAGACAAGCTCCTCCCTTGAGA

Qy 1080 CCTGTGGATTTGAAA--AGATACTATTTTATTATTATTATTGTGACAAAAATGT---T
              |||||||||||||||| |||||||||||||||||||||||||||||||||
Db 1224 CCTGTGGATTTTTAAAACAGATATTATTTTTATTATTTGTGACAAAAATGTTGAT

Qy 1135 GGATATTAAAGAGAATAAATCA 1156
              |||||||||| |||||||||
Db 1284 GGATATTAAATAGAATAAGTCA 1305

RESULT 10
BD062758
LOCUS   BD062758   1373 bp   DNA   linear   PAT 27
DEFINITION A tumor necrosis factor related ligand.
ACCESSION   BD062758
VERSION   BD062758.1   GI:22608361
KEYWORDS   JP 2001505407-A/2.
SOURCE   unidentified
  ORGANISM unidentified
            unclassified.
REFERENCE   1 (bases 1 to 1373)
  AUTHORS   Chicheportiche,Y. and Browning,J.L.
  TITLE   A tumor necrosis factor related ligand
  JOURNAL   Patent: JP 2001505407-A 2 24-APR-2001;
            BIOGEN INC, THE FACULTY OF MEDICINE OF THE UNIVERSITY OF GE

NF family related protein
P 200150507-A/2
4-APR-2001
7-AUG-1997 JP 1998508239
7-AUG-1996 US 60/023541,18-OCT-1996 US   60/028515 PR
-1997 US   60/040820
VES CHICHEPORTICHE,JEFFREY L BROWNING
12N15/28,C07K14/525,G01N33/68,C07K16/24,C12N15/11,A61K48/00,
12N5/10,
61K39/395,A61K38/19,C07K14/705,C07K16/24,C12N15/12
trandedness: Double;
opology: Linear;
ey   Location/Qualifiers
DS   1..852.

Location/Qualifiers
1..1373
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

arity   52.6%;   Score 614.6; DB 6;   Length 1373;
onservative 0; Mismatches 219; Indels 70; Gaps   9;

```
TGAGCCTGGGCCTGGCCTGGCTGGCCTCCTGCCTCGTGGTCAGCCTG   61
||||||||||||||||||||||||||||||||||||| ||||||||||
TGGGCCTGGGCCTGGCCTGGCCTCGTGGCCTCTGCCGTGGTCAGTTTG  237

GCTGGGCAACGCTGTCTGCCCAGGAGCCTTCTCGAGGAGGAGTGACAGCAGAGGAC  121
||||||||||||||||||||||||||||||    |||||||||
GCCGGGCATCGCTGTCCGCCCAGGAGCCGTGTCCGCCAGGAGCAGAGGAGCTGGTGGCAGAGGAG  297

GGGAGCCCCCTGAACTGAATCCCAGACAGAGCAGAAAGCCAGGATGTGGTACCTTTC  181
||||||||||||| ||||||||||||||||||||||||||||||||||||||||
AGGAGCCCGTCGGAACTGAATCCCCAGACAGAGAAGAAAGCCAGGATCCTGCGCCTTTC  357

AACAACTAGTCCGGCCTCCGAAGAAGTGCTCTCAAAGGCCGGAAGGCCGGCCTGC  241
||||||||||||||||||||||||||||||| |||||||||| | || ||||||
ACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGAAAACACGGGCTCGA  417

CTATTGCAGCCCATTATGAGGTTCATCCTCGGCCAGGACAGGATGGAGCACAAGCA  301
||||||||||||||||||||||||||||||| |||||||||| || | ||||||
CGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACGGAGCGCCAGGCA  477

TGGGATGGGACAGTGAGTGGCTGGGAAGAGACCAAAATCAACAGTCCAGCCCTCTG  361
|||||||||||||||||||||||||||||||||| | || ||||||||||
TGGACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAGCCCTCTG  537

ACGACCCCAGATTGGGGAATTACAGTCATCACGGCTGGGCTTCTACTACCTGTAC  421
||||||||||||||||||||||| ||||| ||||||||||||||||||||||
ACAACCGGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTACTACCTGTAC  597

AGGTGCACTTTGATGAGGGAAAGCCTGTCTACCTGGACTTGCTGGTGAAC  481
||||||||||||||||||||||||||||||||||||||||||||||||||
AGGTGCACTTTGATGAGGGAAGGCTGTCTACCTGGACTTGCTGGTGGAT  657

TGCTGGCCCTGCGCTGCCTGGAAGAATTCTCAGCCACAGCAGCAAGCTCTCCTGGG  541
||||||||||||||||||||||||||||||||||||||||||| ||||||||||
TGCTGGCCCTGCGCTGCCTGGAGCGAATTCTCAGCCACTGCGGCCGAGTCCCTGGGG  717

AGCTCCGTTGTGTGCCAGGTGTCTGGGCTGTGCCGCTGCGGCGCCAGGGTCTTCCCTT  601
||||||||||||||||||||| ||||||||||||||||||||||||
AGCTCCGCCTCTGCCAGGTGTCTGGGCTGTTGCCCTGGGCCAGGGTCCTCCCTG  777

TCCGCACCCTCCCCTGGGCTCATCTTAAGGCTGCCCCCTTCCTAACCTACTTTGA  661
||||||||||||||||||||||||||||||||||||||||||||||
TCCGCACCCTCCCTGGGCCCATCCCCATCTCAAGGCTGCCCCTTCCTCACCTACTTCGGA  837

TTCAAGTTCACTGAGGGGCCTTGCTTCTCCAGATTCCTTAAACTTTCCCTGGCTCC  721
|||||||||||||||||||||| ||||| |||||||||||||||||||
TCCAGGTTCACTGAGGGGCCCCTGTCTCCCCAGCGTGCCCGGCTCC  897
```

RESULT 11
AR140407
LOCUS   AR140407   1236 bp   DNA   linear   PAT 16
DEFINITION   Sequence 1 from patent US 6207642.
ACCESSION   AR140407
VERSION   AR140407.1   GI:14482903
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE   1 (bases 1 to 1236)
  AUTHORS   Wiley,S.R.
  TITLE   Member of the TNF family useful for treatment and diagnosi
             disease
  JOURNAL   Patent: US 6207642-A 1 27-MAR-2001;
FEATURES   Location/Qualifiers
     source   1..1236
                     /organism="unknown"
                     /mol_type="unassigned DNA"

ORIGIN

Query Match   51.2%;   Score 597.8;   DB 6;   Length 1236;
Best Local Similarity 75.2%;   Pred. No. 3e-147;
Matches   879; Conservative   0; Mismatches 222;   Indels   68;   G

```
GTGCTGGCGCTCGGCGGCCTGGCGCCGCTGGCCTGCCTGGCCTCGGTCGTGGTCGTCAG   2
| ||||||||||| ||||||||||||||||||| |||||||||| |||||||||||||
GCGCTGGCGCTCGGCGGCCTGGCGCCGCTGGCCTCGGTCGTGGTCGTCAG  73

GGGAGCTGGGCAACGCTGTCTGCCCAGGAGCCTTCTCAGGAGGAGCTGACAGCAGA   62
||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
GGGAGACCGGGCCATCGCTGTCCGCCCAGGAGCCTGCGCCCAGGAGGAGCTGGTGGCAGA  133

CGCCGGGAGCCCCTGAACTGAATCCCGAAGAGAAGTGCCTCCTAAAGGCCGGAAGGCGGCC   122
|||||||||||||| |||||||||||||||||||| |||||||||||||||||||||
GACCAGGACCCGTCGGAACTGAATCCCCAGACAGACAGAAGAAAGCCAGGATCCTGCGCC  193

TTGGAACAACTAGTCCGGCCTCGGAAGAAGTGCCTCGAAGAAGTGCCTCCTAAAGGCCGGAAG   182
|||||||||||||||||||||||||||||||||||||||||||
CTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGAAAACACGGGC   253
```

```
AGGAGCAGTCACCACCACCTCCCTACCCCACCCCCACTCCTCCACCCCCTC-GCTGCT  722
|||||            |||||||||||| ||||||||||||||||
-----CCTCGACAGCTCTCGGCAGCTCCTCC--TCAAAGGCAGCCAGCAGAGCTTGTTCACATGTTCCAT  898

GGTCCAGTCCTGTCTCC--TCAAAGGCAGCCAGCAGAGCTTGTTCACATGTTCCAT  781
||||            |||||   |||||||||||||||||||||||||
GCTCCAGACCTGCCTCCCTCCCCTCTCTAGAGGCTGCTGGGCCTGTTCAGTGTTTTCCA  953

------ACAGACGTATCCTGTCTTCTTCTTAACATCCATCCACCACAACTATCCA  838
|||       | ||||||||||||||| |||||||||||||||||||
ACATAAATACACAGTATCCCACTCTTATCTTACAACTCCCCCACTCCACCTCCA  1013

ACTAGCTCCCAAAGCCCCTAC----------------TTATCCTGACTCCCCACC  892
||| |||            |||||||||||||||||||||||
ACTAGCTCCCCAATCCCCGACCCTTTGAGGCCCCCAGTGATCTGACTCCCCCCTG  1073

CACCCGACCAGTGTTTATTGACTTTGTGCAC------------------------  937
||||       |||||||||||||||
CAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGCAGGATGGGTCCAGAAAG  1133

-----CAGGCACTGAGATGGGCTGGACCTGGTGGCAGGAAGCCAGAGAACCTGGGA  969
|||       ||||||||| |||| | ||||||||||||||||| ||||||||||
CACTTCAGGCACTAAGAGGGGCTGGACCTGGCCGGGCCAGGAAGCCAAAGAGACTGGGC  1193

GCCAGAAGTTCCCAACTGTGAGGGGGAAGAGCTGGGACAAGCTCCTCCCTGGA--  1024
||||                  |||||||||||||| |||||||| |||||
GCCAGGAGTTCCCAAATGTGAGGGGCGAGA-AACAAGACAAGCTCCTCCCTTGAGA  1253

CCTGTGGATTTTGAAA--AGATACTATTTTATTATTATTGTGACAAAATGTTAA  1080
||||                    ||||||||||||||||||||||||||||||
CCTGTGGATTTTTAAACAGATATTATTTTATTTTTATTATTGTGACAAAATGTTGA  1312
```

```
KEYWORDS
SOURCE       JP 200151626-A/1.
ORGANISM     Homo sapiens (human)
             Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 1236)
AUTHORS      Wiley,S.R.
TITLE        Member of the tnf family useful for treatment and diagnosi
JOURNAL      Patent: JP 200151626-A 1 04-SEP-2001;
             ABBOTT LABORATORIES
COMMENT      PN   JP 200151626-A/1
             PD   04-SEP-2001
             PF   12-FEB-1998 JP 1998535077
             PR   12-FEB-1997 US   08/798692,10-FEB-1998 US      09/0217(
             STEVEN R WILEY
             PC   C12Q1/68,C12N15/63,C07K14/525,C07K16/28,G01N33/50,G01I
             PC   A61K38/19
             CC   Strandedness: Single;
             CC   Topology: Linear;
             FH   Key             Location/Qualifiers
FEATURES              Location/Qualifiers
     source      1..1236
                 /organism="Homo sapiens"
                 /mol_type="genomic DNA"
                 /db_xref="taxon:9606"
ORIGIN
```

```
Query Match         51.2%;    Score 597.8;  DB 6;    Length 1236;
Best Local Similarity  75.2%;    Pred. No. 3e-147;
Matches  879;  Conservative  0;  Mismatches  222;  Indels  68;  G.
```

```
Qy    2 GTGCTGAGCCTGGGCCTGGCGCGCTGGCCTCCTGCCTCGTGCCTGGTCAG(
          ||||| || ||||||||||||| |||||| |||||||||||||||||||||
Db   73 GCGCTGGGCCTGGCCTGGCGCCTGGCGCGCTGGCCTCGTGCCTGGTGGTCAG'

Qy   62 GGGAGCTGGGCAACGCTGTCTGCCAGGAGCCTTCTCAGGAGGAGCTGACGACAGAGA(
          ||||||||||| ||||||  |||||||||||||||||||||||||||||||||||||
Db  133 GGGAGCCGGGCCATCGCTGTCGCCCAGGAGCCTGCCCAGACAGAGGAGCTGGTGGCAGA(

Qy  122 CGCCCGGAGCCCCCCGTCGGCCTCGAAGAGTGCTCCTAAAGCCGGAAGGGCGGGCC(
          |||||||| |||||||||||||||||||||||||||||||||| |||||||||||||
Db  193 GACCAGGACCCCGTCGGCCTGGGCCTCGAAGAGTGCTCCTAAAGCCGGAAAACACGGGC'

Qy  182 TTGAACAACTAGTCCGGCCTCGAAGAAGTGCTCCTAAAGGCCGGAAGGGCGGGCC'
          ||||| |||||||||||||||||||||||||||||||||| ||||||||||||||
Db  253 CTGACCGACTAGTTCGGCCTCGAAGAAGTGCACCTAAAGCCGGAAAACACGGGC'

Qy  242 CGAGCTATTGCAGCCCATTATGAGGTTCATCCTCGGCCACCAGGATGAGCACA(
          |||||| ||||||||||||||||||||||||||||| |||||||||| |||||
Db  313 AGAGCGATCGCAGCCCATTATGAGGTTCATCCACGACCTGGACAGGACCGGAGCGGCCA(

Qy  302 GGTGTGGATGGACAGTGAGTGCTGGAAGACCAAAATCAACAGCTCCAGCCC'
          |||||||| |||| ||||||||||||||||||||| ||||||||||||||||
Db  373 GGTGTGGACGGGACCAGTGAGTGGCTGGAGGAAGCCAGATCAACAGCTCCAGCCC'

Qy  362 CGCTACGACACCGCCAGATTGGGAATTTACAGTCATCAGGGCTGGGCTCTACTACC1
          |||||||||||||| |||||||||||||||| |||| |||||||||||||||||||
Db  433 CGCTACAACCGCCAGATCGGGAGTGGGAGTTTATAGTCACCCGGGCTGGGCTCTACTACC1

Qy  422 TGTCAGGTGCCACTTTGATGAGGGAGAAAGGCTGTCTACCTGAAGCTGGACTTGCTGGT(
          |||||||||||||||||||||||||||||||||| ||| |||||||| ||||||||||
Db  493 TGTCAGGTGCACTTTGATGAGGGAGAAAGGCTGTCTACCTGAAGCTGGACTTGCTGGT(

Qy  482 GGTGTGCTGGCCTGCGCTGGGCGCTGGCCTGGAAGAATTCTCAGCCACCAGCAAGCTCCC'
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  553 GGTGTGCTGGCCCTGCGCCTGGCCTGCCTGGAGGAATTCTCAGCCACTGCGGCGAGTTCCCT'

Qy  542 CCCCAGCTCCGTTTGTGCCAGGTGCTCTGGGCGTGTTGCCGCTGCGCGCCAGGGTCTTC(
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  613 CCCCAGCTCCGCCTCTGCCTCTCTCGGCCCAGTGTCTCGGCCGTGTTGGGCCGTGTTCCGCGCCAGGGTCCTC(

Qy  602 CGGATCCGCCACCCCCCGTCGGGCTCATCTTAAGGCTGCGCCTTCCTAACCTACTT'
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
(sequences continued)

TATTGCAGCCATTATGAGGTTCATCCTCGGCCAGGACAGGAGGATGGAGCACAAGCA 301
||||| || ||||||||||||||||||||||||||||||||||||||||||
GATCGCAGCCCATTATGAAGTTCATCCACCAGCGACCTGGACAGGACGGAGCCCAGGCA 372

GGATGGGGACAGTGAGTGGCTGGCTGGGAGAGACCAAAATCAACAGCTCCAGCCCTCTG 361
||||||||||||||  |||| |||||||||||||||||||||||||||||||
GGACGGGGACAGTGAGTGCTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAGCCCTCTG 432

CGACCCGCCAGATTGGGAATTTACAGTCATCAGGGCTCGGCTCTACTACCTGTAC 421
||||||| |||||||||||||||||||||||||||||||||||||||||||||
CAACCGCCAGATCGGGGGAGTTTATAGTCACCCGGGCTCGGCTCTACTACCTGTAC 492

GGTGCACTTGATGATGAGGGAAAGGCTGTCTACCTGACTTGCTGGTGAAC 481
|||||||| || ||| |||||||||||||||||||||||||||
GGTGCACTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGGAT 552

GCTGCCCTGCGCTGCTGGAAGAATTCTCAGCCACAGCAGCAAGCTCCTCGG 541
||||||||||||||||||||||||||||||||||||||||||||||||||
GCTGCCCTGCGCTGCTGGAGGAATTCTCAGCCACTGCGGCGAGTTCCCTCGGG 612

GCTCCGTTTGTGCCAGGTGTCTGGGCTGTTGCCGCTGCGCGCCAGGGTCTTCCCTT 601
|||||| ||||||||||||||||||||||||||||||||||||||||||
GCTCCGCCCTGCCAGGTGTCTGGGCTGTGTTGGGCCGTGTTCCGCGCCAGGGTCTCCCTG 672

CCGGCACCCTCCCCCTGGGCTCATCTTAAGGCTGCCCCTTCCTAACCTACTTTGA 661
|||||||||||||||||||||||||||||||||||||||||||||||||
CCGCACCCTCCCCCTGGGCCCATCTCAAGGCTGCCCCTTCCTAAACCTACTTCCGGA 732

TCAAGTTCACTGAGGGGCCGGCCTTGCTCTCCCAGATTCCTTAAACTTCCCTGGCCTCC 721
||||||||||||||||||||||  |||||||||
CCAGGTTCACTGAGGGGCCCCGGTTCTCCCCGCAGTGTCTCCAGCTGCCGGCTCC 792

CATCACCACACCTCCCCTACCCCACCCCACTCCTCCACCCCCTC-GCTGCTCCTT 780
|| |||  || ||||||||
CCTCGACAGCTCTCTGGGCACCCGGTCCCCTCCTGCCGCCCTCAGCGCGCTCTTT 847

AGTCCTCGTCTCTCC--TCAAAGGCAGCCAGAGCTTGTTCACATGTTTCCATTCC- 837
|||| || || |||||||||||||||||||||||||
AGACCTGCCCCTCCCCTCTCTAGAGGCTGCCTGGGCCGTCTGTTTTTCCATCCC 907

-ACAGAACGTATCCTTGCTCTTCTTAAACATCCCATCCCACACAACTATCCACCTC 891
||||||||||||||||||||||||||||||||||||
AATACAGTATTCCCACTCTTCTATCTTACACAACCCACCGCCACTCTCCACCTC 967

CCTCCCCAAAGCCCCTAC---------TTATCCCTGACTCCCCCCACT 936
|||||||||| |||||||||
CTCCCCAATCCCCTGAGGCCCCCAGTGATCTCGACTCCCCCCTGGCCA 1027

GACCACGTGTTTATTGACTTTGTGCAC-------- 968
||||||| ||  ||
CCCCAGGGCATTGTGTTCACTGTACTCTCTGGGCAAGGATGGTTCCAGAGAGACCC 1087

CAGGCACTGAGATGGCGGGCTGGACCTGGTGCCAGGAAGCCAGAGAACCTGGGACTAG 1023
||||||||||||||||||||||||||||||||||||||||||||||
CAGGCACTAAGAGGGGCTGGACCTGGCGCCAGGAAGCCAGGAAGCACTGGGCCTAG 1147

AAGTTCCCAACTGTGAGGGGGGAAGAGCTCGGGACAAGCTCCTCCTGGA----TC 1079
||||||| |||
GAGTGTCCCAAATGTGAGGGGGCGAGA-AACAAGACAAGCTCCTCCCTTGAGAATTC 1206

GGATTTTGAAAAGATACTATTTTT 1108
|||||| ||||||||||||||||
GGATTTTTAAAAACAGATATTATTT 1235

GGTGTGCTGGCCTGCGCTGGGCGCTGGCCTGGAAGAATTCTCAGCCACCAGCAAGCTCCC'
GGTGTGCTGGCCCTGCGCCTGGCCTGCCTGGAGGAATTCTCAGCCACTGCGGCGAGTTCCCT'

CCCCAGCTCCGTTTGTGCCAGGTGCTCTGGGCGTGTTGCCGCTGCGCGCCAGGGTCTTC(
CCCCAGCTCCGCCTCTGCCTCTCTCGGCCCAGTGTCTCGGCCGTGTTGGGCCGTGTTCCGCGCCAGGGTCCTC(

CGGATCCGCCACCCCCCGTCGGGCTCATCTTAAGGCTGCGCCTTCCTAACCTACTT'
```

```
24       1236 bp    DNA      linear    PAT 27-AUG-2002
of the tnf family useful for treatment and diagnosis of
24
24.1  GI:22602730
```

JOURNAL

Submitted (06-JUN-2003) NIH Intramural Sequencing Center,
Grovemont Circle, Gaithersburg, MD 20877, USA
On Jun 6, 2003 this sequence version replaced gi:27753660.

COMMENT

```
----------- Genome Center
    Center: NIH Intramural Sequencing Center
    Center code: NISC
    Web site: http://www.nisc.nih.gov
    Contact: nisc_zoo@nhgri.nih.gov
----------- Project Information
    Center project name: dcf
    Center clone name: 258K06
-----------
```

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicat
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8X average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associate
with a Phrap-derived quality score.

```
----------- Summary Statistics
    Sequencing vector: plasmid; n/a; 100% of reads
    Chemistry: Dye-terminator Big Dye; 100% of reads
    Assembly program: Phrap; version 0.990319
    Consensus quality: 128169 bases at least Q40
    Consensus quality: 128674 bases at least Q30
    Consensus quality: 128972 bases at least Q20
    Insert size: 150000; agarose-fp
    Insert size: 129154; sum-of-contigs
    Quality coverage: 10.78x in Q20 bases; agarose-fp
    Quality coverage: 12.52x in Q20 bases; sum-of-contig:
-----------
```

```
* NOTE: This is a 'working draft' sequence. It currently
*   consists of 12 contigs. Gaps between the contigs
*   are represented as runs of N. The order of the pieces
*   is believed to be correct as given, however the sizes
*   of the gaps between them are based on estimates that hav
*   provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
*       10521: contig of 10521 bp in length
*       10522: gap of unknown length
*       10621: gap of unknown length
*       13327: contig of 2706 bp in length
*       13328: gap of unknown length
*       13427: gap of unknown length
*       28924: contig of 15497 bp in length
*       28925: gap of unknown length
*       29024: gap of unknown length
*       29025: contig of 10177 bp in length
*       39201: gap of unknown length
*       39202: gap of unknown length
*       39301: contig of 2605 bp in length
*       41906: gap of unknown length
*       41907: gap of unknown length
*       42006: contig of 28089 bp in length
*       70095: gap of unknown length
*       70096: gap of unknown length
*       70195: contig of 7366 bp in length
*       70196: gap of unknown length
*       77561: contig of 16500 bp in length
*       77562: gap of unknown length
*       77661: gap of unknown length
*       94161: contig of 3721 bp in length
*       94162: gap of unknown length
*       94261: gap of unknown length
*       97982: contig of 7318 bp in length
*       97983: gap of unknown length
*       98082: gap of unknown length
*       98083: contig of 22355 bp in length
*      105400: gap of unknown length
*      105401: gap of unknown length
*      105500: contig of 2299 bp in length
*      105501: gap of unknown length
*      127855: gap of unknown length
*      127856: gap of unknown length
*      127955: contig of 2299 bp in length
*      127956: contig
```

FEATURES

source   1..130254

Location/Qualifiers
1..130254
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="Brown Norway"

TCCGCACCTCCCTGGGCCCATCTCAAGGCTGCGCCCCTTCCTCACCTACTTCGGA   732
||||||||||||||||||||||||||
TTCAAGTTCACTGAGGGGCCCTTGCTCTCCAGATTCCTTAAACTTTCCTGGCTCC   721
|||||||||||||||||||
TCCAGGTTCACTGAGGGGCCCTGCTCTCCCCGCAGTCGTCCCAGGCTGCCGGCTCC   792

GCATCACCACACTCCCTACCCCACCCCCACTCCTCCACCCCCTC-GCTGCTCCTT   780
|| ||| ||||| |||||||||||||||||||||||
-CCTGCAGCTCTGGGCACCCGGTCCCCTCTGCCCCTTCAGCCGCTCTTT   847

CAGTCCTGCTCTCC--TCAAAGGCAGCCAGAGCTTGTTCACATGTTTCCATTCC-   837
||| ||||| |||||||||||||||
CAGACCTGCCCCTCCCCTCTAGAGGCGCTGCCTGGGCCTGTTCACGTGTTTTCCATCCC   907

--ACAGACGTATCCTTGCTCTCTTCTTAACATCCCATCCACCAAACTATCCACCTC   891
|||||||||||||||||||||
AAATACAGTATTCCACTCTTATCTTACAACAACCCACTCTCCACCCTC   967

GCTCCCAAAGCCCCTAC--------------TTATCCCTGACTCCCCACCCCACT   936
|| || ||||||||||
GCTCCCCAATCCCTGACCCCTTTGAGGCCCCAGTGATCTCGACTCCCCCCTGGCCA   1027

CGACCACGTGTTTATTGACTTTGTGCAC-----------------------   968
||| || ||||||||||
CCCCCAGGGGCATTGTTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCC   1087

-CAGGCACTGAGATGGGCTGGACCTGGTGGCAGGAAGCCAGAGAAACCTGGGACTAG   1023
|||||||||||||||||||
TCAGGCACTAAGAGAGGGCTGGACCTGGCCGCAGGAAGCCAAAGAGACTGGGCCTAG   1147

GAAGTTCCCAACTGTGAGGGGAAGAGCTGGGACAAGCTCCTCCCCTGGA----TC   1079
|||||||||||||||||||
GGAGTTCCCAAATGTGAGGGGCGAGA-AACAAGACAAGCTCCTCCCTTGAGAATTC   1206

TTGGATTTTGAAAAGATACTATTTTT   1108
||||||||||||||||||
TTGGATTTTTAAAACAGATATTATTT   1235

---

95   130254 bp   DNA   linear   HTG   06-JUN-2003
norvegicus clone RP31-258K6 strain Brown Norway, WORKING
SEQUENCE, 12 ordered pieces.
95
95.3   GI:31442440
TGS_PHASE2; HTGS_DRAFT.
norvegicus (Norway rat)
norvegicus
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

ses 1 to 130254)
llis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
ley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Cariaga,K.,
, Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X.,
J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
B., Idol,J.R., Karlins,E., Kwong,P., Laric,P., Lee-lin,S.-Q.,
i,R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Masiello,C.,
i,B., McDowell,J., Paguirigan,C., Pearson,R., Portnoy,M.E.,
A., Reddix-Dugue,N., Schandler,K., Schueler,M.G., Shah,K.,
C., Stantripop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V.,
J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
omparative Sequencing Initiative
ished
ses 1 to 130254)
E.D.
Submission
ted (30-OCT-2002) NIH Intramural Sequencing Center, 8717
ont Circle, Gaithersburg, MD 20877, USA
ses 1 to 130254)
E.D.
Submission

```
/db_xref="taxon:10116"
/clone="RP31-258K6"
/clone_lib="RP31"
1..10521
/note="assembly_fragment
clone_end:SP6
vector_side:left"
10622..13327
13428..28924
/note="assembly_fragment"
29025..39201
/note="assembly_fragment"
39302..41906
/note="assembly_fragment"
42007..70095
/note="assembly_fragment"
70196..77561
/note="assembly_fragment"
77662..94161
/note="assembly_fragment"
94262..97982
/note="assembly_fragment"
98083..105400
/note="assembly_fragment"
105501..127855
/note="assembly_fragment"
117804..130254
/note="clone overlaps with GenBank Accession Number
AC134317 clone RP31-421B16 (center project name dce)"
127956..130254
/note="assembly_fragment
missing T7 clone end on 3' end of insert"
```

```
rity      48.5%;   Score 566.6;  DB 2;   Length 130254;
          89.1%;   Pred. No. 1.1e-118;
nservative      0;  Mismatches   74;   Indels    8;  Gaps    5;

GGACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTCGTGAACGGT 484
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCACTTTGATGAGGGAAGGCAGTCTACCTGAAGCTGGACTTGCTGGTGAATGT 72516

GGCCCTGCGCTGCCTGGAAGAATTCTCAGCCACAGCAGCAAGCTCTCCTGGGCCC 544
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGCCCTGCGCTGCCTGGAAGAATTCTCAGCCACAGCAGCGAGCTCTCCTGGGCCC 72576

CCGTTTGTGCCAGTGTCTGGGCGTGTTGCCGTCGGCCCAGGGTCTTCCCTTCGG 604
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCGTCGTGCCAGTGTCTGGGCGTGTTGCCTCTCGCGGCCCAGGGTCTTCCCTTCGG 72636

CACCCTCCCCTGGGCTCATCTTAAGGCTGCCCCCTTCCTAACCTACTTTGGACTC 664
||||||||||||||||||||||||||||||||||||||||||||||||||||||
TACCCTCCCCTGGGCTCATCTTAAGGCTGCCCCCTTCCTTACCCTACTTTGGACTC 72696

AGTTCACTGAGGGCCTTGCTCTCCCAGATTCCTTAAACTTTCCC----TGGCTC 720
||||||||||||||||||||||||||||||||||||||| |||||    ||||||
AGTTCACTGAGGGGGTCCTGCTCTCTCCCAGGTTCTTTAAACTTTAACCTAACAGCTC 72756

GCATCACCACCACCTCCCTACCCCACCCCCACTCTCCTCACCCCCCTGCTGCTCCTT 780
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCACCACAGCACCTCCCCACCGCCCCACTCTCCTCCACCCTCTAGCTGTTCCTT 72816

AGTCCCTGT-CTCTCCCTCAAAGGCAGCCAGAGCTTGTTCACATG-TTTCCATTCCA 838
|||||||| ||||||||||||||||||||||||||||||||||||| ||||||||||
AGTCCTGTCTCCTCTCTTCCAAGGCAGCCAGGGCTTGATGACATGTTTTCCATTCCA 72876

GTATCCTTGCTCTTCT-TAACATCCCATCCACCACCACTATCCACCTCACTAGC 897
||||||||||||||||  |||||||||||||||||||||||||||||||||||
CATATTCCCGTTCTTCTGTAACCTCCCAACCCCCACCCCACCAGTGCCCGCCTCACGAGC 72936

AAGCCCTACTTATCCCTGCACTCCCCCACCACTCACCCGACCACGTGTTTATTG 957
||||| |||||  |||||   ||||| ||  |||  |||| |||  |||||
```

```
Db   72937  TCCCAAAGCCCCACGTA-CCGGACTACCCGCCCCAACCCTGGGACCCTGTTTT
             ||||||||||||||||| ||||||||||||||||||||||||||||||||
Qy     958  ACTTTGTGCACCAGGCACTGAGATGGGCTGACCTGGTGGTCAGAGAAGCCAGAGAAC
             ||| ||||| ||||||||||||||||||||||||||| ||||||||||||||
Db   72996  ACTATGTGCATCAGGCATCAAGATAGGCTGGACCTGGTCGGTCAGGAAGCCAAAGAACC
             |||||||||||||||||||||||||||||||||||||| |||||||||||||
Qy    1018  GACTAGGCCAGAAGTTCCCAACTGTGAGGGGGAAGAGCTGGGAGAGCAAGCTCCTCCC
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   73056  GACTAGGCCAGAAGTTCCCAAATGAGGGAAGAGATGAGAGACAAGCTCCTCCC
             |||||||||||||||| |||||||||||||||||||||||||||||||||||
Qy    1078  TCCCTGTGGATTTTGAAAAGATACTATTTTATTATTGTGACAAAATGTTAAA
             |||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   73116  TCCCTGTGGATTTTGAAAAGATACTATTTTTATTATTGTGACAAAATGTTAAA
             ||||||||||||||||||||||||||| ||||||||||||||||
Qy    1138  TATTAAAGAGAATAAATCATGATTTCTCTTC 1168
             |||||||||||||||||||||||||||||
Db   73176  TATTAAAAGAGAATAAATCATGATTTCTCTTC 73206
```

RESULT 14
AC119115/c
LOCUS   AC119115   165316 bp   DNA   linear   HTG 19-
DEFINITION Rattus norvegicus clone CH230-320N23, WORKING DRAFT SEQUENC
            unordered pieces.
ACCESSION   AC119115
VERSION   AC119115.4   GI:25100662
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE   Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleo
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murina
            Rattus.
REFERENCE   1 (bases 1 to 165316)
AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alde
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,A., Baca,M., Baden,H.,
            Baldwin,D., Bandaranaike.D., Barber,M., Barnstead,M., Benal
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,K., Calderon,E.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Sou
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederi
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., E
            Egan,A., Escotto,M. Eugene,C., Evans,C.A., Falls,T., Fan,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., F
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Ga
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Gu
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Ja
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jo
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,A., Lopez,J.,
            Lorensuhewa,L., Louiseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mang
            Mangum,B., Mapua,P., Martin,R., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore
            Morgan,M., Morris,M., Morris,S., Munidasa,M., Murphy,M., N.
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Par
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., R
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Rigg
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
```

e.M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
., Wang,S., Warren,J., Warren,R., Wei.X., White,F.,
ms,G., Willson,R., Wieczyk,R., Wooden,H., Worley,K.,
,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
hausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
ock,G. and Gibbs,R.A.
ished

ses 1 to 165316)
,K.C.

Submission
ted (25-APR-2002) Human Genome Sequencing Center, Department
ecular and Human Genetics, Baylor College of Medicine, One
Plaza, Houston, TX 77030, USA

ses 1 to 165316)
nome Sequencing Consortium.

Submission
ted (19-NOV-2002) Human Genome Sequencing Center, Department
ecular and Human Genetics, Baylor College of Medicine, One
Plaza, Houston, TX 77030, USA.

19, 2002 this sequence version replaced gi:23616728.
quence in this assembly is a combination of BAC based reads
ole genome shotgun sequencing reads assembled using Atlas
//www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
feature table below represents a scaffold in the Atlas
ly (a 'contig-scaffold'). Within each contig-scaffold,
dual sequence contigs are ordered and oriented, and separated
ed gaps filled with Ns to the estimated size. The sequence
tend beyond the ends of the clone and there may be sequence
s within a contig-scaffold that consist entirely of whole
shotgun sequence reads. Both end sequences and whole genome
n sequence only contigs will be indicated in the feature

------ Genome Center
nter: Baylor College of Medicine
nter code: BCM

b site: http://www.hgsc.bcm.tmc.edu/
ntact: hgsc-help@bcm.tmc.edu
------ Project Information
nter project name: GUXG
nter clone name: CH230-320N23
------ Summary Statistics
ssembly program: Phrap; version 0.990329
nsensus quality: 155971 bases at least Q40
nsensus quality: 157321 bases at least Q30
nsensus quality: 158253 bases at least Q20
stimated insert size: 159662; sum-of-contigs estimation
ality coverage: 7x in Q20 bases; sum-of-contigs estimation

: Estimated insert size may differ from sequence length.
ee http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
: This is a 'working draft' sequence. It currently
ists of 5 contigs. The true order of the pieces
not known and their order in this sequence record is
trary. Gaps between the contigs are represented as
s of N, but the exact sizes of the gaps are unknown.
record will be updated with the finished sequence
oon as it is available and the accession number will
reserved.

1	159511:	contig of 159511 bp in length
59512	159611:	gap of unknown length
59612	160669:	contig of 1058 bp in length
50670	160769:	gap of unknown length
50770	161864:	contig of 1095 bp in length
51865	161964:	gap of unknown length
51965	163701:	contig of 1737 bp in length
53702	163801:	gap of unknown length
53802	165316:	contig of 1515 bp in length.

Location/Qualifiers
1 ..165316

---

misc_feature
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-320N23"
1..2103
/note="wgs end extension
clone end:T7"
8259...9143
/note="clone_boundary
clone end:T7
site:
end_sequence:BZ171094"
46341..159014
/note="clone_boundary
clone end:Sp6
site:
end_sequence:BZ171096"

misc_feature

misc_feature

ORIGIN

Query Match   48.5%; Score 566.6; DB 2; Length 165316;
Best Local Similarity 89.1%; Pred. No. 1.1e-138;
Matches 669; Conservative 0; Mismatches 74; Indels 8; C

```
Qy    425 CAGGTGCACTTGATGAGGGAAAGGCTGCTACCTGAAGCTGGACTTGCTGGTGA
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 111973 CAGGTGCACTTGATGAGGGGAAGGCAGTCTACCTGAAGCTGGACTTGCTGGTGA

Qy    485 GTGCTGGCCCTGCGCTGCCTGGAAGAATTCTCAGCGCCAGCAGCAAGCTCTCTG
              ||||||||||||||| ||||||| |||||||| |||||||||||||||||||||
Db 111913 GTGCTGGCCCTGCGCTGCCTGGAAGAATTCTCAGCCACAGCAGCAGCTCTCCTG

Qy    545 CAGCTCCGTTGTGCCAGTGCTGGCCAGGTGTCCGGCGTGTTGCCGCCAGGGTCTTCCC
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 111853 CAGCTCCGTCGTGTGCCAGTGCTGTCGGGCTGTTGCCTCGGGCCAGGGTCTTCCC

Qy    605 ATCCGCACCCTCCCCTGGGCTCATCTTAAGGCTGCCCCCTTCCTAACCTACTTTG
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 111793 ATCCGTACCCTCCCCTGGGCTCATCTTAAGGCTGCCCCCTTCCTTACCTACTTTG

Qy    665 TTTCAAGTTCACTGAGGGGCCTTGCTCTCCAGATTCCTTAAACTTTCCC----T
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db 111733 TTTCAAGTTCACTGAGGGGTCCTGCTCTCTCCAGGTTCTTTAAACTTTAACCTAAC

Qy    721 CAGGAGCATCACCACACTCCCTACCCCACCTCCTCCACCCCCCGCTGC
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db 111673 CCAGAGCACCACAGCACCTCCCCACCCCACCTCCTCCACCCCTCTAGCTGT

Qy    781 GGTCCAGTCCTGT-CTCTCCTCAAAGGCAGCAGCAGAGCTTGTTCACATG-TTTCCA
              ||||||||||||| |||||||||||||||||||||||||||||||||||
Db 111613 GCTCCAGTCCTGTCTCTCTCCTCCAAGGCAGCAGCAGGGCTTGATGACATGTTTCCA

Qy    839 CAGACGTATCCTTGCTCTTCT-TAACATCCCATCCCACCACAACTATCCACCTCA
              ||||| ||||||||||||||| |||||||||||||||||||||||||||||||||
Db 111553 CAGACATATTCCCGTTCTTCTGTAACCTCCCACCCCACCACAGCTGCCCGCCCTCA

Qy    898 TCCCAAAGCCCCTACTTATCCCTGACTCCCCCCACCCACCCGACCACGTGTT
              |||||||||||||||||||||||||| |||||||||||||||||||||||
Db 111493 TCCCAAAGCCCCACGTA-CCCGACTACCCCGCCCAACCCTGGGACCCCGTTTT

Qy    958 ACTTTGTGCACCAGGCACTGAGATGGGCTGGACCTGGTGCGCAGGAAGCCAGAGAA
              |||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db 111434 ACTATGTGCATCAGGCATCAAGATAGGCTGGACCTGGTGGCAGGAAGCCAAAGAA

Qy   1018 GACTAGGCCAGAGTTCCCAACTGTGAGGGGGAAGAGCTGGGGACAAGCTCCTCC
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 111374 GACTAGGCCAGAGTTCCCAAATGTGAGGGGGAAAGAGAGATGGAGACAAGCTCCTCC

Qy   1078 TCCCTGTGGATTTTGAAAAGATACTATTTTATTATTGTGACAAAATGTTAA
              ||||||| ||||||||||| |||||||||| |||| |||||||||||
Db 111314 TCCCTGTGGATTTTGAAAAGATACTATTTTTATTATTTTATTGTGACAAAATGTTAA

Qy   1138 TATTAAAGAGAATAAAATCATGATTTCTCTTC 1168
              ||||||||||||||||||||||||||||||||
```

AAGAGAATAAATCATGATTCTCTTC 111224

3   223877 bp   DNA   linear   HTG 10-MAY-2003
norvegicus clone CH230-154B15, WORKING DRAFT SEQUENCE, 3
ed pieces
3.8 GI:30521223
GS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
norvegicus (Norway rat)
norvegicus
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

es 1 to 223877)
.Marie. Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
echi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
us,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
nd,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,M.,
, Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
lez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
orgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
ne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
lez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
y,S., Kelly,S., Kelly,K., Khan,Z., King,L., Kovar,C.,
., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,J., Liu,J.,
Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
newa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
ari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
ey,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
rljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
ris,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
emeh,O., Okwuonu,G., Olarinpunagoon,A., Pal,S., Parks,K.,
nak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
C.F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,
.,, Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
.B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
.C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
s,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
A., Sodergren,E., Song,X.-Z., Sorelle,M., Sosa,J.,
a,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
.T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
ns,G., Willson,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
hausen,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
ock,G. and Gibbs,R.A.
   Submission
es 1 to 223877)
ished
.K.C.
   Submission
ted (06-NOV-2001) Human Genome Sequencing Center, Department
ecular and Human Genetics, Baylor College of Medicine, One

REFERENCE
AUTHORS
TITLE
JOURNAL

Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 223877)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Dep
of Molecular and Human Genetics, Baylor College of Medicin
Baylor Plaza, Houston, TX 77030, USA

COMMENT
On May 10, 2003 this sequence version replaced gi:25008075
The sequence in this assembly is a combination of BAC base
and whole genome shotgun sequencing reads assembled using
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig d
in the feature table below represents a scaffold in the At
assembly (a 'contig-scaffold'). Within each contig-scaffol
individual sequence contigs are ordered and oriented, and
by sized gaps filled with Ns to the estimated size. The se
may extend beyond the ends of the clone and there may be s
contigs within a contig-scaffold that consist entirely of
genome shotgun sequence reads. Both end sequences and whol
shotgun sequence only contigs will be indicated in the fea
table.
--------   Genome Center
     Center: Baylor College of Medicine
     Center code: BCM
     Web site: http://www.hgsc.bcm.tmc.edu
     Contact: hgsc-help@bcm.tmc.edu
--------   Project Information
     Center project name: GI0K
     Center clone name: CH230-154B15
--------   Summary Statistics
     Assembly program: Atlas 3.0;
     Consensus quality: 214785 bases at least Q40
     Consensus quality: 216908 bases at least Q30
     Consensus quality: 218593 bases at least Q20
     Estimated insert size: 227169; sum-of-contigs estimati
     Quality coverage: 7x in Q20 bases; sum-of-contigs esti
----
* NOTE: Estimated insert size may differ from sequence len
*   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_da
* NOTE: This is a 'working draft' sequence. It currently
*   consists of 3 contigs. The true order of the pieces
*   is not known and their order in this sequence record is
*   arbitrary. Gaps between the contigs are represented as
*   runs of N, but the exact sizes of the gaps are unknown.
*   This record will be updated with the finished sequence
*   as soon as it is available and the accession number will
*   be preserved.
*   1   221327: contig of 221327 bp in length
*   221328   221427: gap of unknown length
*   221428   222652: contig of 1225 bp in length
*   222653   222752: gap of unknown length
*   222753   223877: contig of 1125 bp in length.

FEATURES   Location/Qualifiers
     source   1..223877
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-154B15"
     misc_feature   1..1822
                     /note="wgs_contig"
     misc_feature   complement(217607..218056)
                     /note="clone_boundary
                     clone_end:T7
                     site:EcoRI
                     end_sequence:RWBBO08TJB"

ORIGIN

Query Match   48.5%;   Score 566.6;   DB 2;   Length 223877;
Best Local Similarity 89.1%; Pred. No. 1.1e-138;
Matches 669; Conservative   0; Mismatches   74;   Indels   8;   G

Qy   425 CAGGTGCACTTTGATGAGGGAAAGGCTCTCACCTGAAGCTGGACTTGCTGTGAA
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 147649 CAGGTGCACTTTGATGAGGGAGAGGGAAGGCAGTCTACCTGAAGCTGGACTTGCTGTGAA

```
CTGGCCCTGCGCTGCCTGCTGGAGAGAATTCTCAGCCACAGCAGCAAGCTCTCCTGGGCCC  544
||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
CTGGCCCTGCGCTGCCTGCTGGAGAGAATTCTCAGCCACAGCAGCGAGCTCTCCTGGGCCC  147768

CTCCGTTGTGCCAGGTGTCTGGGCTGTTGCCGCTGGGCCAGGGTCTTCCCTTCGG  604
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTCCGTCGTGTGCCAGGTGTCTGGGCTGTTGCCTCTGCGGCCAGGGTCTTCCCTTCCG  147828

CGCACCCTCCCCTGGGCTCATCTTAAGGCTGCCCCCTTCCTAACCTACTTTGGACTC  664
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
CGTACCCTCCCCTGGGCTCATCTTAAGGCTGCCCCCCTTCCTTACCTACTTTGGACTC  147888

CAAGTTCACTGAGGGGCCTTGCTCTCCCAGATTCCTAAACTTTCCC---TGGCTC  720
||||||||||||||||||||||||||||||||||| |||||||||||  |||||||
CAAGTTCACTGAGGGGTCCTGCTCTCCCAGGTTCTTAAACTTTAACCTAACAGCTC  147948

GAGCATCACCACACCTCCCTACCCCACCCCACTCCTCCACCCCCTGCTGCTCCTT  780
|||||||||||||||| ||||||||||| ||||||||||||||||| |||||||||
GAGCACCACAGCACCTCCCCACCCCACGCCCACTCCTCCACCCCTAGCTGTTCCTT  148008

CCAGTCCTGT-CTCTCCCAAAGGCAGCCAGAGCTTGTTCACATG-TTTCCATTCCA  838
|||||||||| |||||||||||||||||||||||||||||| |  |||||||||||
CCAGTCCTGTCCTCTCCTCTTCCAAGGCAGCCAGGGCTTGATGACATGTTTTCCATTCCA  148068

ACGTATCCTTGCTCTTCT-TAACATCCCATCCACCACAACTATCCACCTCACTAGC  897
| ||||||||||||||||| ||| ||||| || ||||||| |||||||||  |||
ACATATTCCCCGTTCTTCTGTAACCTTCCCACCCCACCACCACAGCTGCCCGCCTCACGAGC  148128

CAAAGCCCCTACTTATCCTGACTCCCCCACCACTCACCCGACCACGTGTTATTG  957
||||||||||||||| |||| ||| ||| ||||||| |||||| || |||||||
CAAAGCCCCACCGTA-CCCGACTACCCCGCCCCGCCCAACCCTGGGACCCTGTTTTTACTG  148187

TTGTGCACCAGGCACTGAGATGGGCTGGACCTGGTGGCAGGAAGCCAGAGAAACCTGG  1017
||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
ATGTGCATCAGGCATCAAGATAGGCTGGACCTGGTGGCAGGAAGCCAAAGAAACCTGG  148247

CAGGGCAGAAGTTCCCAACTTGTGAGGGGGAAGAGCTGGGACAAGAGCTCCTCCCCTGGA  1077
||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
CAGGGCCAGAAGTTCCCAAATGTGAGGGGAAGAGATGGGACGACAAGCCTCCTCCCCTGGA  148307

CTGTGGATTTTGAAAAGATACTATTTTTATTATTATTATTGTGACAAAATGTTAAAATGGA  1137
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTGTGGATTTTGAAAAGATACTATTTTTATTATTATTATTGTGACAAAATGTTAAAATGGA  148367

CAAAGAGAATAAAATCATGATTTCTCTTC  1168
|||||||||||||||||||||||||||||
CAAAGAGAATAAAATCATGATTTCTCTTC  148398
```

April 8, 2004, 02:01:42
5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

search, using sw model

11 7, 2004, 17:30:19 ; Search time 2831.52 Seconds
(without alignments)
12318.149 Million cell updates/sec

09-245-198A-1

gtgctgagcctggcctgg........ataaatcatgattctcttc 1168

NTITY_NUC
op 10.0 , Gapext 1.0

13289 seqs, 1493109276 residues

s satisfying chosen parameters: 55026578

th: 0
th: 2000000000

nimum Match 0%
ximum Match 100%
sting first 45 summaries

T:*
em_estba:*
em_esthum:*
em_estin:*
em_estmu:*
em_estov:*
em_estpl:*
em_estro:*
em_htc:*
gb_est1:*
gb_est2:*
gb_htc:*
gb_est3:*
gb_est4:*
gb_est5:*
em_estfun:*
em_estom:*
em_gss_hum:*
em_gss_inv:*
em_gss_pln:*
em_gss_vrt:*
em_gss_fun:*
em_gss_mam:*
em_gss_mus:*
em_gss_pro:*
em_gss_rod:*
em_gss_phg:*
em_gss_vrl:*
gb_gss1:*
gb_gss2:*

the number of results predicted by chance to have a
than or equal to the score of the result being printed,
d by analysis of the total score distribution.

SUMMARIES

ry ch	Length	DB	ID	Description
.7	1033	11	AK020909	AK020909 Mus muscu
.6	892	14	CB204861	CB204861 AGENCOURT
.6	918	10	BF577781	BF577781 602092080
.8	665	13	BY742288	BY742288 BY742288

5		549	47.0	549	AK044387	2237	11	AK044387	Mu
6		533.4	45.7	533.4	CG565104	543	29	CG565104	OS
7		519.4	44.5	519.4	BI871711	731	12	BI871711	60
8		510	43.7	510	CG29394	554	29	CG29394	OS
9		507.4	43.4	507.4	BI870393	728	12	BI870393	60
10		504	43.2	504	AW763237	561	10	AW763237	ur
11		497	42.6	497	BE628951	533	10	BE628951	uu
12		489.6	41.9	489.6	BG404836	650	12	BG404836	60
13	c	488.8	41.8	488.8	BQ208433	687	13	BQ208433	UI
14		481.2	41.2	481.2	AW917574	584	10	AW917574	ES
15		480.8	41.2	480.8	CB849011	939	14	CB849011	MR
16		480.4	41.2	480.4	BI819200	777	12	BI819200	60
17		476	40.8	476	CG653257	482	29	CG653257	OS
18		473	40.5	473	BI766766	834	12	BI766766	60
19		467.8	40.1	467.8	BQ707185	948	13	BQ707185	AG
20		464.8	39.8	464.8	CF126539	666	14	CF126539	UI
21		456.4	39.1	456.4	BI596681	828	12	BI596681	60
22		448.2	38.4	448.2	CG584545	621	29	CG584545	OS
23		440.6	37.7	440.6	CB141389	545	14	CB141389	K-
24		437.8	37.5	437.8	AA221610	471	9	AA221610	my1
25	c	430.2	36.8	430.2	CG554711	498	29	CG554711	OS
26		426.8	36.5	426.8	BU759448	624	13	BU759448	UI
27		422.6	36.2	422.6	CF126932	697	14	CF126932	UI
28		421.6	36.1	421.6	CG609156	474	29	CG609156	OS
29	c	421.6	36.1	421.6	BQ671259	963	13	BQ671259	AC
30		420	36.0	420	CG646715	483	29	CG646715	OS
31		419.2	35.9	419.2	BQ884231	940	13	BQ884231	AC
32	c	410.4	35.1	410.4	BX634398	456	13	BX634398	BX
33		406.8	34.8	406.8	CG568080	471	29	CG568080	OS
34		406.4	34.8	406.4	BM921213	1071	12	BM921213	AC
35		400.4	34.3	400.4	CG611020	418	29	CG611020	OS
36		393.2	33.7	393.2	CG596702	494	29	CG596702	OS
37		390.6	33.4	390.6	AA870722	445	9	AA870722	vg2
38		390.2	33.4	390.2	CG498076	493	29	CG498076	OS
39		386.6	33.1	386.6	CG590009	531	29	CG590009	OS
40		383.2	32.8	383.2	CG573612	468	29	CG573612	OS
41		366.6	31.4	366.6	BQ674188	951	13	BQ674188	AC
42		360	30.8	360	BE654876	360	10	BE654876	UI
43		353.2	30.2	353.2	CG525153	483	29	CG525153	OS
44	c	353	30.2	353	AI854476	405	9	AI854476	UI-
45		350.6	30.0	350.6	CG599845	414	29	CG599845	OS

ALIGNMENTS

RESULT 1
AK020909

LOCUS   AK020909   1033 bp   mRNA   linear   HTC 20-
DEFINITION Mus musculus adult retina cDNA, RIKEN full-length enriched
            clone:A930030D13 product:tumor necrosis factor (ligand)
            superfamily, member 12., full insert sequence.
ACCESSION   AK020909
VERSION   AK020909.1 GI:12861640
KEYWORDS   HTC; CAP trapper.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleo
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murina
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE   High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki
  TITLE   Normalization and subtraction of cap-trapper-selected cDNAs
            prepare full-length cDNA libraries for rapid discovery of n
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159

a.,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H. Itoh,M.,
., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
co,R. Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
ke,S., Inoue,K., Togawa,Y., Izawa,K., Ohara,E., Watahiki,M.,
,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
i,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
integrated sequence analysis (RISA) system--384-format
cing pipeline with 384 multicapillary sequencer
Res. 10 (11), 1757-1771 (2000)
13
61

KEN Genome Exploration Research Group Phase II Team and the
Consortium.
onal annotation of a full-length mouse cDNA collection
409, 685-690 (2001)

NTOM Consortium and the RIKEN Genome Exploration Research
Phase I & II Team.
is of the mouse transcriptome based on functional annotation
770 full-length cDNAs
420, 563-573 (2002)
ses 1 to 1033)

.J., Aizawa,K., Akahira,S., Akimura,T., Arai.A., Aono,H.,
a.,T., Bono.H., Carninci,P., Fukuda,S., Fukunishi,Y.,
,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
wa,T., Kato,K., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
a,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
,Y., Toya,T., Yanamura,T., Yasunishi,A., Yoshida,K.,
o,M., Muramatsu,M. and Hayashizaki,Y.
Submission
ted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of
al and Chemical Research (RIKEN), Laboratory for Genome
opedia Project of Genome Exploration Research Group in Riken
c Sciences Center and Genome Science Laboratory in RIKEN.
ation Research Group, RIKEN Genomic Sciences Center (GSC),
Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
wa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
tp://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
-45-503-9216)
visit our web site (http://genome.gsc.riken.go.jp/) for
r details.
library was prepared and sequenced in Mouse Genome
opedia Project of Genome Exploration Research Group in Riken
c Sciences Center and Genome Science Laboratory in RIKEN.
on of Experimental Animal Research in Riken contributed to
e mouse tissues. First strand cDNA was primed with a primer
GAGAGAAGAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTTVN 3'], cDNA was
ed by using trehalose thermo-activated reverse transcriptase
bsequently enriched for full-length by cap-trapper. cDNA went
h two rounds of normalization to Rot = 20.0 and subtraction
= 458.8. Second strand cDNA was prepared with the primer
r of sequence [5'
SAGATTCTGAGTTAATTAAATTAATCCCCCCCCCCC 3']. cDNA was cleaved
amHI and XhoI. Vector: a modified pBluescript KS(+) after
xcision from Lambda FLC I. Cloning sites, 5' end: SalI; 3'
amHI. Host: DH10B. Retina RNA was provided by Stefano
cich (Department of Neurobiology, Harvard Medical School, 220
od Ave. Boston, MA 02115, USA) whose assistance is
ully acknowledged

Location/Qualifiers
1..1033
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"

/db_xref="FANTOM_DB:A930030D13"
/db_xref="MGI:1911867"
/db_xref="taxon:10090"
/clone="A930030D13"
/tissue_type="retina"
/clone_lib="RIKEN full-length enriched mouse cDNA

CDS   /dev_stage="adult"
<1..433
/note="unnamed protein product; putative
tumor necrosis factor (ligand) superfamily, membe
(MGD|MGI:1916833, GB|NM_023517, evidence: BLASTN,
match=787)"
/codon_start=2
/protein_id="BAB32249.1"
/db_xref="GI:12861641"
/translation="IAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSS
GEFTVIRAGLYYLYCQVHFDEGKAVYLKiDLLVNGVLALRCLEEFSATA
LCQVSGLLPLRPGGSLRIRTLPWAHLKAAPFLTYFGLFQVH"

ORIGIN

Query Match   75.7%;   Score 883.8; DB 11;   Length 1033;
Best Local Similarity 99.4%;   Pred. No. 1.9e-219;
Matches 918; Conservative   0; Mismatches   3;   Indels   3;   G

Qy   248 ATTGCAGCCCATTATGAGGTTCATCCTCGGCCAGGACAGGAGATGGAGCACAAGCAGG
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2 ATTGCAGCCCATTATGAGGTTCATCCTCGGCCAGGACAGGAGATGGAGCACAAGCAGG

Qy   308 GATGGGACAGTGAGTGGCTGGGAAGAGACCAAAATCAACAGCTCCAGCCCCTGCG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62 GATGGGACAGTGAGTGGCTGGGAAGAGACCAAAATCAACAGCTCCAGCCCCTGCG

Qy   368 GACCGCCAGATGGGAATTTACAGTCATCAGGGCTGGGCTCTACTACCTGTACTG
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   122 GACCGCCAGATTGGGAATTTACAGTCATCAGGGCTGGGCTCTACTACCTGTACTG

Qy   428 GTGCACTTTGATGAGGGAAAGGCTGTCTACCTGAACTGGACTTGCTGGTGAACGG
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   182 GTGCACTTTGATGAGGGAAAGGCTGTCTACCTGAACTGGACTTGCTGGTGAACGG

Qy   488 CTTGGCCCTGCGCTGCCTGGAAGAATTCTCAGCCAGCCACAGCAAGCTCCTGGGCC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   242 CTGGCCCTGCGCTGCCTGGAAGAATTCTCAGCCAGCCACAGCAAGCTCTCGGGCC

Qy   548 CTCCGTTGTGCCAGGTGTCTGGGCTTGCCGCTCGGGCCAGGGTCTTCCCTTCG
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   302 CTCCGTTGTGCCAGGTGTCTGGGCTTGCCGCTCGGGCCAGGGTCTTCCCTTCG

Qy   608 CGCACCCTCCCTGGGCTCATCTTAAGGCTGCCCCCTTCCTAACCTACTTTGGACT
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   362 CGCACCCTCCCTGGGCTCATCTTAAGGCTGCCCCC-TCCTAACCTACTTTGGACT

Qy   668 CAAGTTCACTGAGGGGCCTTGCTCTCCCAGATTCCTAAACTTCCCTGGCTCCAG
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   422 CAAGTTCACTGAGGGGCCTTGCTCTCCCAGATTCCTAAACTTCCCTGGCTCCAG

Qy   728 ATCACCACACCTCCCTACCCCACCTCCTCCACCCCCTGCTGCTCCTTGGT
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   482 ATCACCACACCINTCTACCCCACCTCCTCCACCCCCTAGCTGCTCCTTGGT

Qy   788 TCCTGT-CTCTCCTCAAAGGCAGCCAGAGCTTGTTCACATG-TTTCATTCCACAG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   542 TCCTGTCCTCTCTCAAAGGCAGCCAGAGCTTGTTCACATGTTTTCATTCCACAG

Qy   846 ATCCTTGCTCTTC-TTAACATCCCATCCACCAACTATCCACTCACTCAGTCTCC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   602 ATCCTTGCTCTTTTTAACATCCCATCCACCAACTATCCACTCACTCAGTCTCC

Qy   905 GCCCCTACTTATCCCTGACTCCCCGACTCCCCGACCACCACGTGTTATTGACT
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   662 GCCCCTACTTATCCCTGACTCCCCGACTCCCCGACCACCACGTGTTATTGACT

Qy   965 GCACCAGGCACTGAGATGGGCTGGACCTGGTGGCAGGAAGGCCAGAGAACCTGGAC

```
Db     61  GACCGCCAGATGGGGAATTACAGTCATCAGGGCTGGGCTCTACTACCTGTACTGT
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    781  AGGCACTGAGATGGGGCTGGACCTGGTGCGGCAGGAAGCCAGAGAACCTGGGACTAGG   781

Qy    428  GTGCACTTTGATGAGGGAAAGCTGTCTACCTGACTTGCTGCTGGTGAACGGT
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  GTGCACTTTGATGAGGGAAAGCTGTCTACCTGACTTGCTGCTGGTGAACGGT   1084

Qy    488  CTGGCCCTGGCTGCCTGGAAGAATTCTCAGCCACAGCAGCAAGCTCTCCTGGGCCC
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  CTGGCCCTGGCGTCGCTGGAAGAATTCTCAGCCACAGCAGCAAGCTCTCCTGGGCCC   841

Qy    548  CTCCGTTTGTGCCAGGTGTCTGGGCTGTTGCCGCTGCGCGCCAGGGTCTTCCCTTCGG
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  CTCCGTTTGTGCCAGGTGTCTGGGCTGTTGCCGCTGCGCGCCAGGGTCTTCCCTTCGG   1144

Db    608  CGCACCCCCCCCTGGGCTCATCTTAAGGCTGCCCCCTTCCTAACCTACTTGGACTC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  CGCACCCTCCCCTGGGCTCATCTTAAGGCTGCCCCCTTCCTAACCTACTTGGACTC

Qy    668  CAAGTTCACTGAGGGGCCTTGCTCTCCCAGATTCCTTAAACTTTCCCTGGCTCCAGG
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  CAAGTTCACTGAGGGGCCTTGCTCTCCCAGATTCCTTAAACTTTCCCTGGCTCCAGG

Qy    728  ATCACCACACCTCCTACCCCACCCCACTCCTCCACCCCCTGCTGCTCCTTGGTC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  ATCACCACACCTCCTACCCCACCCCACTCCTCCACCCCCTGCTGCTCCTTGGTC

Qy    788  TCCTGT-CTCTCCTCCAAAGGCAGCCAGAGCTTGTTCACATG-TTTCCATTCCACGA
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  TCCTGTCCTCTCCTCCAAAGGCAGCCAGAGCTTGTTCACATGTTTCCATTCCACAGA

Qy    846  ATCCTTGCTCTTC-TTAACATCCCATCCAACAACTATCCACTCACTAGCTCCC
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  ATCCTTGCTCTTCTTTAACATCCCATCCAACAACTATCCACTCACTAGCTCCC

Qy    905  GCCCCTACTTATCCTGACTCCCCCACCCATCACCCGACCACGTGTTTATTGACTT
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  GCCCCTACTTATCCTGACTCCCNCACCCACCCGAACACGTGTTTATTGACTT

Qy    965  GCACCAGGCACTGAGATGGGCTGGACCTGGTGCGGCAGGAGAACCTGGACGGAC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  GACCANGCACTGAGATGGGCTGGACCTGGTGCGGCAGGAGAACCTGGACGGAC

Qy   1025  CCAGAAGTTCCCAACTGTGAGGGGG-AAGAGCTGGGGACAAGCT--CCTCCCTGGAT
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  CCAGAAGTTCCCAACTGTGAGGGGGAAAGAGCTGGGGACCAAGCTCCCTCCCTGGAT

Qy   1083  G--TGGATTTTGAAAAGATACTATTTTT---ATTATTATTGTGACAAATGTTAAA
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  GGTGGAATTTTGAAAAAGAACNAATTTTGTAATTATATATTGGTGACAAATGTTTAA

Qy   1138  TAT  1140
           |||
Db    841  GAT  843
```

RESULT 3

```
BF577781
LOCUS         BF577781               918 bp    mRNA         linear       EST 12-
DEFINITION    602092080F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:42(
              mRNA sequence.
ACCESSION     BF577781
VERSION       BF577781.1  GI:11651493
KEYWORDS      EST.
SOURCE        Mus musculus (house mouse)
ORGANISM      Mus musculus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele(
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murina
REFERENCE     1  (bases 1 to 918)
AUTHORS       NIH-MGC http://mgc.nci.nih.gov/,
TITLE         National Institutes of Health, Mammalian Gene Collection (f
JOURNAL       Unpublished (1999)
COMMENT       Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
```

Procurement: Jeffrey E. Green, M.D.
Library Preparation: Life Technologies, Inc.
Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
equencing by: Incyte Genomics, Inc.
distribution: MGC clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL at:
/image.llnl.gov

LLAM9767 row: p column: 20
uality sequence start: 17
uality sequence stop: 724.
Location/Qualifiers
1..918
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4206595"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Co24"
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

Qy   682 GGCCT 686
                |||||
Db   658 GGCCT 662

RESULT 4
BY742288
LOCUS   BY742288   665 bp mRNA   linear   EST 17
DEFINITION   BY742288 RIKEN full-length enriched, adult retina Mus musc
             clone A930030D13 5', mRNA sequence.
ACCESSION   BY742288
VERSION   BY742288.1 GI:2716686
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murir
REFERENCE   1 (bases 1 to 665)
AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kc
             Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
             Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
             Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,
             Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsu
             Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
             Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.
             Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
             Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
             Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kar
             Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konaga
             Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.
             Maltais,L., Marchionni,L., McKenzie,L., Miki,H. Nagashima
             Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
             Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandrar
             Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringw
             Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,
             Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomit
             Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.
             Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Y
             Yang,I., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,
             Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
             Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizaw
             Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K.,
             Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shil
             Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Land
             Rogers,J., Birney,E. and Hayashizaki,Y.
TITLE   Analysis of the mouse transcriptome based on functional ar
             of 60,770 full-length cDNAs
JOURNAL   Nature 420, 563-573 (2002)
MEDLINE   22354683
PUBMED   12466851
COMMENT   Contact: Yoshihide Hayashizaki
             Laboratory for Genome Exploration Research Group, RIKEN Ge
             Sciences Center(GSC), Yokohama Institute
             The Institute of Physical and Chemical Research (RIKEN)
             1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-004
             Tel: 81-45-503-9222
             Fax: 81-45-503-9216
             Email: genome-res@gsc.riken.go.jp,
             URL:http://genome.gsc.riken.go.jp/
             Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,
             Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F
             Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima
             Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakam
             Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Saka
             Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagam
             Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashiz
             Direct Submission
             Computational Analysis of Full-Length Mouse cDNAs Compare
             Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
             Normalization and subtraction of cap-trapper-selected cDN
             prepare full-length cDNA libraries for rapid discovery of
             genes. Genome Res. 10 (10), 1617-1630 (2000)
             RIKEN integrated sequence analysis (RISA) system--384-for

arity   53.6%;   Score 625.8; DB 10; Length 918;
onservative 99.2%; Pred. No. 4.7e-152;
              Mismatches 2;   Indels 3;   Gaps   3;

GCCTGCCTGGCCTCCTGCTCGTGTGTCGTGGGGAGCTGGGCAACGCTGTC 81
||||||||||||||||||||||||||||||||||||||||||||||||||
GCCTGCCTGGCCTCCTGCTCGTGTGTCGTGGGGAGCTGGGCAACGCTGTC 60

CAGGAGCCTTCTCAGGAGGAGCTGACAGCAGAGGACCGCCGGGAGCCCCCTGAACT 141
||||||||||||||||||||||||||||||||||||||||||||||||||
CAGGAGCCTTCTCAGGAGGAGCTGACAGCAGAGGACCGCCGGGAGCCCCCTGAACT 120

CCCCAGACAGAGGAGAAAGCCAGGATGTGGTACCTTCTTGGAACAACTAGTCCGGCC 201
||||||||||||||||||||||||||||||||||||||||||||||||||
CCCCAGACAGAGGAGAAAGCCAGGATGTGGTACCTTCTTGGAACAACTAGTCCGGCC 180

AGAAGTGCTCCTAAAGGCCGGAAGGCCGGCCTTGCGCCAGCTATTGCAGCCCCATTA 261
||||||||||||||||||||||||||||||||||||||||||||||||||
AGAAGTGCTCCTAAAGGCCGGAAGGCGCGG-CTCGCCAGCTATTGCAGCCCCATTA 239

GTTCATCCTCGGCCAGGACAGGATGGAGCACAAGCAGGTGTGGATGGGACAGTGAG 321
||||||||||||||||||||||||||||||||||||||||||||||||||
GTTCATCCTCGGCCAGGACAGGATGGAGCACAAGCAGGTGTGGATGGGACAGTGAG 299

TGGGAAGAGACCAAAATCAACAGCTCCAGCCCCTGCGCTACGACCGCCAGATTGG 381
||||||||||||||||||||||||||||||||||||||||||||||||||
TGGGAAGAGACCAAAATCAACAGCTCCAGCCCCTGCGCTACGACCGCCAGATTGG 359

TTTACAGTCATCAGGGCTGGCCTCTACTACCTGTACTGTCAGGTGCACTTTGATGA 441
||||||||||||||||||||||||||||||||||||||||||||||||||
TTTACAGTCATCAGGGCTGGCCTCTACTACCTGTACTGTCAGGTGCACTTTGATGA 419

AAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGAACCGGTGTGCTGGCCTGCGCTG 501
||||||||||||||||||||||||||||||||||||||||||||||||||
AAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGAACCGGTGTGCTGGCCTGCGCTG 479

GAAGAATTCTCAGCCACAGCAGCAAGCTCTCCTGGGGCCCCAGCTTCCGTTGTGCCA 561
||||||||||||||||||||||||||||||||||||||||||||||||||
GAAGAATTCTCAGCCACAGCAGCAAGCTCTCCTGGGGCCCCAGCTCCGTTGTGCCA 539

TCTGGGGCTGTTGCCGCTGCGGCCAGGGTCTTCCCTTCCCATCCGACCCCTCCCCTG 621
||||||||||||||||||||||||||||||||||||||||||||||||||
TCTGGGGCTGTTGCCGCTGCGGCCA-GGTCTTCCCTTCCCATCCGACCCCTCCCCTG 598

CATCTTAAGGCTGCCCCCTTCCTAACTACTTTGGACTCTTTCAAGTTCACTGAGG 681
||||||||||||||||||||||||||||||||||||||||||||||||||
CATCTTAAGGCTG-CCCCTTCCTAACTAACTTTGGACTCTTTCAAGTTCACTGAGG 657

ing pipeline with 384 multicapillary sequencer. Genome Res.
, 1757-1771 (2000)
er-based methods for the mouse full-length cDNA
pedia: real-time sequence clustering for construction of a
ndant cDNA library. Genome Res. 11 (2), 281-289 (2001)
ibrary was prepared and sequenced in Mouse Genome
pedia Project of Genome Exploration Research Group in Riken
Sciences Center and Genome Science Laboratory in RIKEN.
n of Experimental Animal Research in Riken contributed to
mouse tissues.
RNA was provided by Dr. Stefano Gustincich (Department of
ology, Harvard Medical School, 220 Longwood Ave., Boston, MA
USA) whose assistance is gratefully acknowledged
visit our web site (http://genome.gsc.riken.go.jp) for
details.

```
Location/Qualifiers
1..665
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="A930030D13"
/tissue_type="retina"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult retina"
/note="Site 1: SalI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGGAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTTTVN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 20.0 and subtraction to Rot = 459.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5'GAGAGAGATTCGAGTAATAAATAATCCCCCCCCCCCC
3'].cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I. ~Retina RNA was provided by Stefano Gustincich,
Department of Neurobiology, Harvard Medical School, 220
Longwood Ave., Boston, MA02115, USA, whose assistance we
gratefully acknowledge."
```

```
   51.8%;      Score 605.6; DB 13;  Length 665;
rity  96.4%;   Pred. No. 7.7e-147;
nservative    0; Mismatches  22;   Indels   2;  Gaps   2;

AGCCCATTATGAGGTTCATCCTCGGCCAGGACAGGATGGAGACAAGCAGGTGTG 307
||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGCCCATTATGAGGTTCATCCTCGGCCAGGACAGGATGGAGACAAGCAGGTGTG 61

GCAGTGAGTGGCTGGGAAGAGACCAAAATCAACAGTTCCAGCCCTTCGCGTAC 367
|||||||||||||||||||||||||||||||||||||||||||||||||||
GACAGTGAGTGGCTGGGAAGAGACCAAAATCAACAGTTCCAGCCCTTCGCGTAC 121

CCAGATTGGGAATTACAGTCATCAGGGCTGGGCTTCTACTACCTGTACTGTCGAG 427
||||||||||||||||||||||||||||||||||||||||||||||||||||
CCAGATTGGGAATTACAGTCATCAGGGCTGGGCTTCTACTACCTGTACTGTCAG 181

CTTTGATGAGGGAAAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGAACGGTGTG 487
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTTTGATGAGGGAAAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGAACGGTGTG 241

CCTTGGCGTCGGCTGGGAAGAATTCTCAGCCACAGCAGCAAGCTCTCCTGGGCCCCAG 547
||||||||||||||||||||||||||||||||||||||||||||||||||||
CCTCGCGTCGGCTGGGAAGAATTCTCAGCCACAGCAGCAAGCTCTCCTGGGCCCCAG 301

TTTTGTGCCAGGTGTCTGGGCTGTTGCCGCTGCGCGCCAGGGTCTTCCCTTCGGATC 607
||||||||||||||||||||||||||||||||||||||||||||||
TTTTGTGCCAGGTGTCTGGGCTGTTGCCGCTGCGCGCCAGGGTCTTCCCTTCGGATC
```

```
Db   302  CTCCGTTGTGCCAGGTGTCTGGGCTGTTGCCGCTGCGGCCAGGGTCTTCCCTTCG
Qy   608  CGCACCCTCCCCTGGGCTCATCTTAAGGCTGCCCCCTTCCTAACCTACTTTGGACT
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   362  CGCACCCTCCCCTGGGCTCATCTTAAGGCTGCCCCCTTCCTAACCTACTTTGGACT
Qy   668  CAAGTCACTGAGGGGCCTTGCTCTCCCAGATTCCTAAACTTTCCTGGCTCCAG
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   422  CAAGTTCACTGAGGGGCCTTGCTCTCCCAGATTCCTAAACTTTCCTGGCTCCAG
Qy   728  ATCACCACCACCTCCTACCCCACCCCCATCCTCCACCCCCTCGCTCCTTGGT
          ||||||||||||| ||||||||||||| |||||||||||||| |||| ||||
Db   482  ATCACCACCACCTNTCTACCCCACCCCTCACTCCTCCACCCCCTAGCTGCTCTTGGT
Qy   788  TCCTGT-CTCTCCTCCAAAGGCAGCCAGACAGAGCTTGTTCACATGTTCATTCCACAGA
          ||||||  ||||||| ||||||||| |||||||||||||||||| |||| ||||||||
Db   542  TCCTGTCCTCTCCTCANAGGCAGACAGAGCTTGTTCACATGNTTTTCATCCCACAGA
Qy   847  TCCTTGCTCTT-CTTAACATCCCATCCACCACAACTATCCACCTCACTAGTCCCA
          |||||||||||  |||||| ||||||||||||||||||||| |||||||||||||
Db   602  TCCTTGCTCTTCCTTACATTCCATCCACCACAACTATTCACCTTACTAGTCCAC
Qy   906  CCC 908
          |||
Db   662  CCC 664
```

RESULT 5
AK044387

```
LOCUS        AK044387               2237 bp    mRNA    linear   HTC 20
DEFINITION   Mus musculus adult retina cDNA, RIKEN full-length enriched
             clone:A930010H24 product:tumor necrosis factor (ligand)
             superfamily, member 12., full insert sequence.
ACCESSION    AK044387
VERSION      AK044387.1  GI:26336423
KEYWORDS     HTC; CAP trapper.
SOURCE       Mus musculus (house mouse)
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murin
REFERENCE    1
  AUTHORS    Carninci,P. and Hayashizaki,Y.
  TITLE      High-efficiency full-length cDNA cloning
  JOURNAL    Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE    99279253
   PUBMED    10349636
REFERENCE    2
  AUTHORS    Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,
             Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizak
  TITLE      Normalization and subtraction of cap-trapper-selected cDNA
             prepare full-length cDNA libraries for rapid discovery of 
  JOURNAL    Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE    20499374
   PUBMED    11042159
REFERENCE    3
  AUTHORS    Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carr
             Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,K., It
             Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Har
             Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashi
             Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watal
             Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., 
             Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizak
  TITLE      RIKEN integrated sequence analysis (RISA) system-384-forma
             sequencing pipeline with 384 multicapillary sequencer
  JOURNAL    Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE    20530913
   PUBMED    11076861
REFERENCE    4
  AUTHORS    The RIKEN Genome Exploration Research Group Phase II Team 
             FANTOM Consortium.
  TITLE      Functional annotation of a full-length mouse cDNA collecti
  JOURNAL    Nature 409, 685-690 (2001)
REFERENCE    5
```

Matches 562; Conservative 0; Mismatches 0; Mismatches 0; Indels 3; C

```
Qy      1 GGTGCTGAGCCTGGGCGCTGGGCTGGCCTGGCCTTGGCCTCCTGGCCTCCTGCCTGCCTGTCTGTGGTC
          |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db    417 GGTGCTGAGCCTGGGCGCTGGGCTGGCCTGGCCTTGGCCTCCTGGCCTCCTGCCTGCCTGTCTGTGGTC

Qy     61 GGGGAGTGGGCAACGTGTCTGCC----CAGGAGCCTTCTCAGGAGGAGGAGCTGACAC
          |||||||||||||||||||||||    |||||||||||||||||||||||||||||||
Db    477 GGGGAGTGGGCAACGTGTCTGCCCACGCAGGAGCCTTCTCAGGAGGAGGAGCTGACAC

Qy    118 GGACCGCCGGGAGCCCCTGAACTGAATCCCAGACAGAGGAAAGCCAGGATGC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    537 GGACCGCCGGGAGCCCCTGAACTGAATCCCAGACAGAGGAAAGCCAGGATGC

Qy    178 TTTCTTGAACAACTAGTCCGGCCTCGAAGAAGTGCTCCTAAAGGCCGGAAGGGGC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    597 TTTCTTGAACAACTAGTCCGGCCTCGAAGAAGTGCTCCTAAAGGCCGGAAGGGGC

Qy    238 TCGCCGAGCTATTGCAGCCCATTATGAGGTTCATCCTCGGCCAGGACAGGATGAC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    657 TCGCCGAGCTATTGCAGCCCATTATGAGGTTCATCCTCGGCCAGGACAGGATGAC

Qy    298 AGCAGGTGTGGATGGGACAGTGAGTGGTGCTGGGAAGAGACCAAAATCAACAGCTCCA
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    717 AGCAGGTGTGGATGGGACAGTGAGTGGTGCTGGGAAGAGACCAAAATCAACAGCTCCA

Qy    358 TCTGCGCTACGACGCCCAGATTGGGAATTACAGTCATCAGGGCTGGGCTCTAC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    777 TCTGCGCTACGACGCCGCCAGATTGGGAATTACAGTCATCAGGGCTGGGCTCTAC

Qy    418 GTACTGTCAGGTGCACTTTGATGAGGGAAAGGCTGTCTACCTGAAGCTGGACTTG
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    837 GTACTGTCAGGTGCACTTTGATGAGGGAAAGGCTGTCTACCTGAAGCTGGACTTG

Qy    478 GAACGGTGTGCTGCCGCCTGGAAGAATTCCAGCCACAGCAGCAAGCC
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    897 GACGGTGTGCTGCCGCCTGGAAGAATTCCAGCCACAGCAGCAAGCC

Qy    538 TGGGCCCCAGCTCCGTTTGTGCCAG  562
          |||||||||||||||||||||||||
Db    957 TGGGCCCCAGCTCCGTTTGTGCCAG  981
```

RESULT 6
CG565104
LOCUS   CG565104   543 bp   DNA   linear   GSS 0:
DEFINITION OST189654 Mus musculus 129Sv/Ev Mus musculus genomic clone
            OST189654, genomic survey sequence.
ACCESSION   CG565104
VERSION   CG565104.1 GI:37351691
KEYWORDS   GSS.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutel
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murii
            1 (bases 1 to 543)
REFERENCE   Zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J.
AUTHORS   Piggott,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Fi
            Friddle,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing
            RPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSSSPLRYDRQIGEFTVIRAG
            Key,B.W. Jr., Kipp,P., Kohlhauff,B., Ma,Z.-Q., Markesich,.
            Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,
            Sparks,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,!
            Zhu,Q., Person,C. and Sands,A.T.
  TITLE   Wnkl kinase deficiency lowers blood pressure in mice: a g(
            screen to identify potential targets for therapeutic inte:
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
COMMENT   Contact: Zambrowicz BP
            OmniBank
            Lexicon Genetics Incorporated
            4000 Research Forest Drive, The Woodlands, TX 77381, USA
            Email: materials@lexgen.com
            Gene trap sequence tag generated by 3' RACE from mouse ES
            described in Zambrowicz et al (Nature. 1998 Apr 9;392(667(

NTOM Consortium and the RIKEN Genome Exploration Research
Phase I & II Team.
is of the mouse transcriptome based on functional annotation
770 full-length cDNAs
420, 563-573 (2002)
ses 1 to 2237)
,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
ida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
, Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kurata,M.,
,, Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
ra,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
i,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
t., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
tsu,M. and Hayashizaki,Y.
Submission
ted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
al and Chemical Research (RIKEN), Laboratory for Genome
ation Research Group, RIKEN Genomic Sciences Center (GSC),
Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
wa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
tp://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
-45-503-9216)
ibrary was prepared and sequenced in Mouse Genome
opedia Project of Genome Exploration Research Group in Riken
c Sciences Center and Genome Science Laboratory in RIKEN.
on of Experimental Animal Research in Riken contributed to
e mouse tissues.
RNA was provided by Dr. Stefano Gustincich (Department of
biology, Harvard Medical School, 220 Longwood Ave., Boston, MA
USA) whose assistance is gratefully acknowledged. Please
our web site for further details.
tp://genome.gsc.riken.go.jp/
tp://fantom.gsc.riken.go.jp/.
Location/Qualifiers
1..2237
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:A930010H24"
/db_xref="MGI:2409696"
/db_xref="taxon:10090"
/clone="A930010H24"
/tissue_type="retina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
346..1578
/note="unnamed protein product; putative
tumor necrosis factor (ligand) superfamily, member 12.
[MGD|MGI:1916833, GB|NM_023517, evidence: BLASTN, 99%,
match=787)"
/codon_start=1
/protein_id="BAC31897.1"
/db_xref="GI:26336424"
/translation="MAARRSQRRRGRRGEPGTALIAPLVLSLGLALACLGLLLVVVSL
GSWATLSAQQEPSQEELTAEDRREPPELNPGTEESQDVVPFLEQLVRPRRSAPKGRKA
RPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSSSPLRYDRQIGEFTVIRAG
LYYLYCQVHFDEGKAVYLKLDLLVNGVLALRCLEEFSATAASSPGPQLRLCQTELQSL
RREVSRLQRSGGPSQKQEARPWQSLWEQSPDVLERWKDGAKSRRRRAVLTQKHKKKHS
VLHLVPVNITSKADSDVTEVMWQPVLRRGRGLEAQGDIVRVWDTGIYLLYSQVLFHDV
TFTMGQVVSREQGRRETLFRCISRMSPSDDPRAYNSCYSAGVFHLHQGDIITVKIPRA
NAKLSLSPHGTFLGFVKL"
2215..2220
/note="putative"
2237
/note="putative"

larity 99.5%; Pred. No. 8.2e-132;   Score 549; DB 11; Length 2237;
         47.0%;

Gene Trap.
Location/Qualifiers
1..543
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129Sv/Ev"
/db_xref="taxon:10090"
/clone="OST189654"
/cell_type="embryonic stem cell"
/clone_lib="Mus musculus 129Sv/Ev"

rity   45.7%;   Score 533.4; DB 29; Length 543;
         98.9%;   Pred. No. 5.2e-128;
nservative   0;   Mismatches   6;   Indels   0;   Gaps   0;

```
         CCTAAAGGCCGGAAGGCGCGGCCTTGCCCAGCTATTGCAGCCATTATGAGGTTC 269
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
         CCTAAAGGCCGGAAGGCGCGGCCTTGCCCAGCTATTGCAGCCATTATGAGGTTC 63

         CGGCCAGGACAGGATGGAGCACAAGCAGGTGTGATGGGACAGTGAGTGGCTGGG 329
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
         CGGCCAGGACAGGATGGAGCACAAGCAGGTGTGATGGGACAGTGAGTGGCTGGG 123

         ACCAAAATCAACAGCTCCAGCCCTCTGCGCTACGACCGCCAGATTGGGAATTTA 389
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
         ACCAAAATCAACAGCTCCAGCCCTCTGCGCTACGACCGCCAGATTGGGAATTTA 183

         ATCAGGGGCTTGGGCTTCTACTACCTGTACTGTCAGGTGCACTTTGATGAGGGAAAGG 449
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
         ATCAGGGGCTTGGGCTTCTACTACCTGTACTGTCAGGTGCACTTTGATGAGGGAAAGG 243

         TACCTGAAGCTGGACTTGCTGGTGAACGGTGTGCTGGGCCCCAGCTCGTTGCCCGCTGCCTCGGAAG 509
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
         TACCTGAAGCTGGACTTGCTGGTGAACGGTGTGCTGGGCCCCAGCTCGTTGCCCGCTGCCTCGGAAG 303

         TCAGCCACAGCAGCAAGCTCTCCTGGGCCCCAGCTCCGTTGTGCCAGGTGTCTG 569
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
         TCAGCCACAGCAGCAAGCTCTCCTGGGCCCCAGCTCCGTTGTGCCAGNGTNTG 363

         TTTGCCGCTGCGGCCCAGGGTCTTCCCTTCGGATCCGCACCCTCCCCTGGGCTCATC 629
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
         TTTGCCGCTGCGGCCCAGGGTCTTCCCTTCNGATCCGCACCCTCCCCTGGGCTCATC 423

         GCTGCCCCCTTCCTAACCTACTTTGGACTTTGGACTTCTTTCAAGTTCACTGAGGGGCCTTGC 689
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
         GNTGCCCCCTTCCTAACCTACTTTGGACTTCTTTCAAGTTCACTGAGGGGCCTTGC 483

         CAGATTCCTTAAACTTTCCCTGGCTCCAGGAGCATCACCACCACCTCCCTACCCCA 749
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
         CAGATTCCTTAAACTTTCCCTGGCTCCAGGAGCATCACCACCACCTCCCTACCCCA 543
```

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information ca
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12034   row: b   column: 07
High quality sequence stop: 728.

FEATURES   Location/Qualifiers
     source   1..731
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:5405478"
                  /tissue_type="adenocarcinoma, cell line"
                  /lab_host="DH10B (phage-resistant)"
                  /clone_lib="NIH_MGC_90"
                  /note="Organ: liver; Vector: pCMV-SPORT6; Site 1:
                  Site 2: SalI; Cloned unidirectionally; oligo-dT p
                  Average insert size 1.7 kb. Library enriched for
                  full-length clones and constructed by Life Techno
                  Note: this is a NIH_MGC Library."

ORIGIN

Query Match   44.5%;   Score 519.4; DB 12; Length 731;
Best Local Similarity 85.3%; Pred. No. 2.7e-124;
Matches 604; Conservative   0; Mismatches 101; Indels   4;   G

```
Qy   16  CCTGGCGCTGGCCTCCTGCCCTCCTGCTGCGTTCGTGGCTCAGCCTCGGGAGCTGGG
         ||||||||||||||||||||||||||||| || |||||||||||||||||||
Db    1  CCTGGCGCTGGCCTCCTGCCCTCCTGCTGCGTTCGTGGCTCAGTTGGGGAGCCGGG

Qy   76  GCTGTCTGCC---CAGGAGCTTCTCAGGAGGAGGAGCTGACAGCAGAGGACCGCCGGG
         |||||||||    ||||||||||||||||||||||||||||||| |||| || ||||
Db   61  GCTGTCGCGCCCAGGAGCAGGAGCCTGCCCAGGAGGAGGAGCTGGTGCGCAGGACCAGG

Qy  133  CCCTGAACTGAATCCCAGACAGAGAAAGCCAGATGGTGTACCTTTCTTGGAAC
         |||||||||||||||||||||||||||| ||||| |||||| ||||||||
Db  121  GTCGGAACTGAATCCCAGACAGAGAAAGCCAGATCCTGCGCCCTTTCCTGAACC

Qy  193  AGTCCGGCTCGAAGAAGTGCTCTAAAGGCCGAAGGCGCGCCTCGGCAGCAGGTGTGG
         ||||||||||||||||||||| |||| || ||| ||||| |||||
Db  181  AGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGAAAAACACGGGCTCGAAGAGCGA

Qy  253  AGCCCATTATGAGGTTCATCCTCGGCCAGGACAGGATGGAGCACAAGCAGGTGTGG
         |||||||||||||||||||||||||||||||| ||| |||||| || || ||
Db  241  AGCCCATTATGAAGTTCATCCAGCATCCACGACGTGGACAGGACAGGACGCAGGTGTGG

Qy  313  GACAGTGAGTGGCTGGGAAGAGACCAAAATCAACAGCTCCAGCCCCTCTGGGCTACG
         |||||||||||||||||||||||||||| ||||| |||||||| ||||| |||||
Db  301  GACAGTGAGTGGCTGGGAGAGAAGCCAGAATCAACAGCTCCAGCCCCTGCGCTACA

Qy  373  CCAGATTGGGGAATTTACAGTCATCAGGGCTGGGCTCTACTACCTGTACTGTCAGG
         ||||||||||||||| ||||| ||||||||||||||||||||||||||||||
Db  361  CCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTACTACCTGTACTGTCAGG

Qy  433  CTTTGATGAGGGAAAGGCTGCTACCTGGACTTGCTGGTGAACGGTGTGC
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  CTTTGATGAGGGAAAGGCTGTCTACCTGGACTTGCTGGTGGATGGTGTGC

Qy  493  CCTGCGCTGCCTGGAAGAATTCTCAGCCACCAGCAGCAGCAAGCTCTCCTGGGCCCCAGC
         |||||||||||||||||||||||| |||||||||||| ||||||||||||||||
Db  481  CCTGCGCTGCCTGGAGGAATTCTCAGCCACTCACTGCGGCCAGTTCCCTCGGGATCC

Qy  553  TTTGTGCCAGGTGTCTGGGCTGTGCGGCCTGGCGGCCCCAGGGTCTTCCCTTCGGATCC
         |||| |||||||||| |||||||||||||||||| ||||||| || |||||||||
Db  541  CCTTCGCCAGGTGTTCGGGGCTGTGTTGGCCCTGTTGGCCCGCGGCCA-GGTCCTCCGCGGATCC

Qy  613  CCTCCCCTGGGCTCATCTTAAGGCTGCCCCCTTCCTTAACCTACTTTGGACTCTTTC
         |||||||||||||| ||||||||| ||||||| ||||| |||||||||||||||
Db  600  CCTCCCCTGGGCCCATCTCAAGGCTGCCCCCCTTCCTCCTCACTACTTTCGGACTCTTCC

Qy  673  TCACTGAGGGGCCTTGCTCCCCAGATTCCTTAAACTTTCCCTGGCTCC 721
         |||||| ||||| ||||||
```

1 GI:16045386

1.1

piens (human)
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 731)
http://mgc.nci.nih.gov/.
al Institutes of Health, Mammalian Gene Collection (MGC)
ished (1999)
: Robert Strausberg, Ph.D.
cgapbs-r@mail.nih.gov
Procurement: ATCC
ibrary Preparation: Life Technologies, Inc.

25F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5405478 5',
equence.   731 bp   mRNA   linear   EST 11-OCT-2001

Qy 660 GACTCTTTCAAGTTCACTGAGGGGCCTTGCTCTCCCCAGATTCTTAAACTTTCCC
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 366 GACTCTTTCAAGTTCACTGAGGGGCCTTGCTCTCCCCAGATTCTTAAACTTTCCC

Qy 720 CCAGGAGCATCACCACACTCCCTACCCCACCCCACCCCCTGCTG(
     ||||||||||||||||||||||||||||||||||||||||
Db 426 CCAGGAGCATCACCACACTCCCTACCCCACCCCACCCCCTGGCTG(

Qy 780 TGGTCCAGTCCTGT-CTCTCTCCAAAGGCAGCCAGAGCTTGTTCACATG-TTTCC;
     |||||||||||||| |||||||||||||||||||||||||||||||| |||||
Db 486 TGGTCCAGTCCTGTCCTCTCTCCTCCAAAGGCAGCCAGAGCTTGTTCACATGTTTCC;

Qy 838 ACAGACGTA 846
     |||||||||
Db 546 ACAGACGTA 554

RESULT 9
BI870393
LOCUS   BI870393 728 bp mRNA   linear EST 1
DEFINITION 6033956441F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5405'
     mRNA sequence.
ACCESSION BI870393
VERSION BI870393.1 GI:16044066
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
REFERENCE 1 (bases 1 to 728)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
     Email: cgapbs-r@mail.nih.gov
     Tissue Procurement: ATCC
     cDNA Library Preparation: Life Technologies, Inc.
     cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL
     DNA Sequencing by: Incyte Genomics, Inc.
     Clone distribution: MGC clone distribution information c
     found through the I.M.A.G.E. Consortium/LLNL at:
     http://image.llnl.gov
     Plate: LIAM12034 row: a column: 12
     High quality sequence stop: 728.
FEATURES   Location/Qualifiers
     source   1..728
           /organism="Homo sapiens"
           /mol_type="mRNA"
           /db_xref="taxon:9606"
           /clone="IMAGE:5405459"
           /tissue_type="adenocarcinoma, cell line"
           /lab_host="DH10B (phage-resistant)"
           /clone_lib="NIH_MGC_90"
           /note="Organ: liver; Vector: pCMV-SPORT6; Site 1
           Site 2: SalI; Cloned unidirectionally; oligo-dT]
           Average insert size 1.7 kb. Library enriched for
           full-length clones and constructed by Life Techn
           Note: this is a NIH_MGC Library."

ORIGIN

Query Match 43.4%; Score 507.4; DB 12; Length 728;
Best Local Similarity 86.7%; Pred. No. 3.7e-121;
Matches 594; Conservative 0; Mismatches 86; Indels 5;

Qy 16 CCTGGCGCTGGCCTGCCTTGGCCTCGTCGGTCGTGTGGTCAGCCTGGGGAGCTGG
     |||||||||||||||||||||| ||||||||||||||||||| |||||||||||
Db 1 CCTGGCGCTGGCCTGCCTTGGCCTCCTCGTCGGTCGGTCGTGTGGTCAGTTTGGGGAGCCGG

Qy 76 GCTGTCTGCC---CAGGAGCCTTCTCAGGAGGAGCTGACAGCAGAGGACCGCGG
     |||||| ||| ||||||||||||||||||||||||||||||||||||||
Db 61 GCTGTCCGCCGCCAGCAGGTGCCCCAGCAGGAGCAGGAGCTGGTGGCAGGAGGAGGACCAG

Qy 133 CCCTGGAACTGAATCCCCAGACAGAGGAAAGCCAGGATGTGGTACCTTTCTTGGAA

source   1..561
                    /organism="Mus musculus"
                    /mol_type="mRNA"
                    /strain="129,C57BL/6J,FVB/N"
                    /db_xref="taxon:10090"
                    /clone="IMAGE:3155633"
                    /tissue_type="tumor, gross tissue"
                    /dev_stage="10 months"
                    /lab_host="DH10B"
                    /clone_lib="NCI_CGAP_Mam3"
                    /note="Organ: mammary; Vector: pCMV-SPORT6; Site
                    Site 2: NotI; Cloned unidirectionally. Primer: O
                    Library construced by Life Technologies. Investi
                    providing samples: Lothar Hennighausen/Chu-Xia De
                    Reference for transgenic model: Xu et al., Nature
                    22, 37-43 (1999)."

ORIGIN

Query Match   43.2%;   Score 504; DB 10;   Length 561;
Best Local Similarity 99.8%; Pred. No. 2.5e-120;
Matches 504;   Conservative   0; Mismatches   1; Indels   0; G

Qy   1 GGTGCTGAGCCTGGGCGCTGGCCTCGCCCTTCGCCTCCTCGTCGTGTGGTCA
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57 GGTGCTGAGCCTGGGCGCTGGCCTCGCCCTTCGCCTCCTCGTCGTGTGGTCA

Qy   61 GGGGAGCTGGGCAACGCTGTCTGCCCAGGAGCCTTCTCAGAGGAGGAGCTGACAGCAG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   117 GGGGAGCTGGGCAACGCTGTCTGCCCAGGAGCCTTCTCAGAGGAGGAGCTGACAGCAG

Qy   121 CCGCCGGGGAGCCCCTGAACTGAATCCCAGACAGGAGGAAAGCCAGGATGTGGTAC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   177 CCGCCGGGGAGCCCCTGAACTGAATCCCAGACAGGAGGAAAGCCAGGATGTGGTAC

Qy   181 CTTGGAACAACTAGTCCGGCCTGAAGAAGTGCTCCTAAAGGCCGGAAGGCGGGC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   237 CTTGGAACAACTAGTCCGGCCTGAAGAAGTGCTCCTAAAGGCCGGAAGGCGGGC

Qy   241 CCGAGCTATTGCAGCCCATTATGAGGTTCATCCTCCGGCCAGGACAGGATGAGGAGCAC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   297 CCGAGCTATTGCAGCCCATTATGAGGTTCATCCTCCGGCCAGGACAGGATGAGGAGCAC

Qy   301 AGGTGTGGATGGGACAGTGACTGAGTGGCTGGGAAGAGACCAAAATCAACAGCTCCAGCC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   357 AGGTGTGGATGGGACAGTGACTGAGTGGCTGGGAAGAGACCAAAATCAACAGCTCCAGCC

Qy   361 GCGCTACGACCGCCAGATTGGGGAATTTACAGTCATCAGGGGCTGGGCTCTACTACC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   417 GCGCTACGACCGCCAGATTGGGGAATNTACAGTCATCAGGGGCTGGGCTCTACTACC

Qy   421 CTGTCAGGTGCTGACTTTGATGAGGAGGAAAGGCTGTCTACCTGAAGCTGGACTTGCTGC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   477 CTGTCAGGTGCTGACTTTGATGAGGAGGAAAGGCTGTCTACCTGAAGCTGGACTTGCTGC

Qy   481 CGGTGTGCTGGCCCTGCGCTGCCTG 505
         |||||||||||||||||||||||||
Db   537 CGGTGTGCTGGCCCTGCGCTGCCTG 561

RESULT 11
BE628951
LOCUS   BE628951   533 bp   mRNA   linear   EST 25
DEFINITION uu30c03.y1 Soares_mammary_gland NMLMG Mus musculus cDNA cl
            IMAGE:3373444 5' similar to TR:O54907 O54907 TNF-RELATED W
            INDUCER OF APOPTOSIS ;, mRNA sequence.
ACCESSION   BE628951
VERSION   BE628951.1 GI:9911639
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murin
REFERENCE   1 (bases 1 to 533)

||||||||||||||||||||||||||||||||||||||||||||
AACTGAATCCCAGACAGAGAAGAAACCAGGATCCTGCGCCCTTCCTGAACCGACT 180

GGCCTCGAAGAGTGCTCCTAAAGGCCGGAAGGCGGGAGGCCTCGCGAGCTATTGC 252
||||||||||||||||||||||||||||||||||||||||||||
GGCTCGCAGAAGTGCACCTAAAGGCCGGAAAACACGGGCTCGAAGAGCGATCGC 240

ATTATGAGGTTCATCCTCCGGCCAGGACAGGATGAGCAAAGCAGGTGTGGATGG 312
||||||||||||||||||||||||||||||||||||||||||||
ATTATGAAGTTCATCCACGACCTGACAGGACGACGGAGGCGCAGGCGTGTGGACGG 300

TGAGTGGCTGGGAAGAGACCAAAATCAACAGCTCCAGCCCTGCGCTACGACCG 372
||||||||||||||||||||||||||||||||||||||||||||
TGAGTGGCTGGGAGGAGACCAGAATCAACAGCTCCAGCCCTGCGCTACAACCG 360

ATTGGGGAATTACAGTCATCAGGGCTGGGCTCTACTACCTGTCAGGTGCA 432
||||||||||||||||||||||||||||||||||||||||||||
ATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTACTACCTGTCAGGTGCA 420

GATGAGGAGGAAAGGCTGTCTACCTGAAGCTGACTTGCTGGTGAACGGTGTGCTGGC 492
||||||||||||||||||||||||||||||||||||||||||||
GATGAGGAGGAAGGCTGTCTACCTGAAGCTGACTTGCTGGTGGATGGTGTGCTGGC 480

CGCTGCCTGAGAATTCTCAGCCACAGCAGCAAGCTCTCCTGGGCCCAGTCCG 552
||||||||||||||||||||||||||||||||||||||||||||
GCTCCTGGAGGAATTCTCAGCCACTGCGGCCAGTTCCCTGGGCCCAGTCCG 540

GCCAGGTGTCTGGGCGTGTTGCCGCTGCGCGGCCAGGGTCTTCCCTTCGGATCCGCAC 612
||||||||||||||||||||||||||||||||||||||||||||
GCCAGGTGTCTGGGCGTGTTGGCCCTGGCGGCCA--GGTCCTCCCTGCGGATCCGCAC 599

CCCTGGGCTCATCTTAAGGCTGCCCCCTTCCTAACCTACTTTGGACTCTTTTCAAGT 672
||||||||||||||||||||||||||||||||||||||||||||
CCCTGGGCCCATCCTCAAGGCTG--CCCCTTCCTCACCTACTTCGGACTCTTCCAGT 658

CGAGGGGCCTTGCTCTCCCCAG 697
||||||||||||||||||||||
GGAGGGGGCCCTGGTCTCCCCG 683

37   561 bp   mRNA   linear   EST 04-MAY-2000
9.y1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:3155633 5'
C to TR:O54907 O54907 TNF-RELATED WEAK INDUCER OF APOPTOSIS
A sequence.
37
37.1 GI:7695174

sculus (house mouse)
sculus
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ses 1 to 561)

AP http://www.ncbi.nlm.nih.gov/ncicgap.
al Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Gene Index
ished (1997)
ESTs: ur70d09.x1
t: Robert Strausberg, Ph.D.
cgapbs-r@mail.nih.gov
Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
Library Preparation: Life Technologies, Inc.
Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
equencing by: Washington University Genome Sequencing Center
distribution: NCI-CGAP clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL at:
llnl.gov/image/html/iresources.shtml

58389
imer: -40RP from Gibco
uality sequence stop: 433.
Location/Qualifiers

AP http://www.ncbi.nlm.nih.gov/ncicgap.
al Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Gene Index
ished (1997)
t: Robert Strausberg, Ph.D.
cgapbs-r@mail.nih.gov
lone is available royalty-free through LLNL ; contact the
Consortium (info@image.llnl.gov) for further information.
83048
imer: -40RP from Gibco
uality sequence stop: 437.
　　　Location/Qualifiers
　　　1..533
　　　/organism="Mus musculus"
　　　/mol_type="mRNA"
　　　/db_xref="taxon:10090"
　　　/clone="IMAGE:3373444"
　　　/sex="female (lactating)"
　　　/tissue_type="mammary gland"
　　　/lab_host="DH10B"
　　　/clone_lib="Soares mammary gland NMLMG"
　　　/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
　　　polylinker; 1st strand cDNA was prepared from mammary
　　　gland tissue from a lactating female, and was then primed
　　　with a Not I - oligo(dT) primer. Double-stranded cDNA was
　　　ligated to Eco RI adaptors (Pharmacia), digested with Not
　　　I and cloned into the Not I and Eco RI sites of the
　　　modified pT7T3 vector. Library is normalized. Library
　　　was constructed by Bento Soares and M. Fatima Bonaldo. "

arity　42.6%;　Score 497;　DB 10;　Length 533;
onservative　0;　Mismatches　0;　Indels　3;　Gaps　3;

```
GTGCTGGCCTGCTGCGCTGCTGGAAGAATTCTCAGCCACAGCAGCAAGCTCTCCTGG 540
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTGCTGGCCTGCTGCGCTGCTGGAAGAATTCTCAGCCACAGCAGCAAGCTCTCCTGG  60

CAGCTCCGTTTGTGCCAGGTGTCTGGGCTGTTGCCGCTGCGCGGCCAGGGTCTTCCCT 600
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAGCTCCGTTTGTGCCAGGTGTCTGGGCTGTTGCCGCTGCGCGGCCAGGGTCTTCCCT 120

ATCCGGCCACCCTCCCCTGGGCTCATCTTAAGGCTGCCCCCTTCCTAACCTACTTTGG 660
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ATCCGGCCACCCTCCCCTGGGCTCATCTTAAGGCTGCCCCCTTCCTAACCTACTTTGG 180

TTTCAAGTTCACTGAGGGCCCTTGCTCTCCCAGATTCCTTAAACTTTCCCTGGCTC 720
|||||||||||||||||||||||||||||||||||||||||||||||||||||
TTTCAAGTTCACTGAGGGCCCTTGCTCTCCCAGATTCCTTAAACTTTCCCTGGCTC 240

AGCATCACCACCACCTCCCTACCCCACCCCACCCCCACCCCCACCCCGGTCTTCCCT 780
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGCATCACCACCACCTCCCTACCCCACCCCACCCCCACCCCCACCCCGGTCTTCCCT 300

CAGTCCTGT-CTCTCCCCAAAGGCAGCCAGAGCTTGTTCACATG-TTTCCATTCCA 838
|||||||| |||||||||||||||||||||||||||||||||| |||||||||||
CAGTCCTGTCCTCTCCCCAAAGGCAGCCAGAGCTTGTTCACATGTTTTCCATTCCA 360

CGTATCCTTGCTCTTC-TTAACATCCCATCCCACCACAACTATCCACCTCACTAGC 897
||||||||||||||| ||||||||||||||||||||||||||||||||||||||
CGTATCCTTGCTCTTCTTTAACATCCCATCCCACCACAACTATCCACCTCACTAGC 420

AAAGCCCCTACTTATCCCTGACTCCCCCACCCCACTCACCCGACCACGTGTTTATTG 957
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
AAAGCCCCTACTTATCCCTGACTCCCCCACCCCACTCACCCGACCACGTGTTTATTG 480

TGTGCACCAGGCACTGAGATGGGCTGGACCTGGTGGCAGGAAGCCAGAG 1010
|||||||||||||||||||||||||||||||||||||||||||||||||
TGTGCACCAGGCACTGAGATGGGCTGGACCTGGTGGCAGGAAGCCAGAG  533
```

BG404836
LOCUS　　BG404836　602420160F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:45270　650 bp　mRNA　linear　EST 12
DEFINITION　mRNA sequence.
ACCESSION　BG404836
VERSION　BG404836.1　GI:13298284
KEYWORDS　EST.
SOURCE　Mus musculus (house mouse)
ORGANISM　Mus musculus
　　　Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
　　　Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murin
REFERENCE　1　(bases 1 to 650)
AUTHORS　NIH-MGC http://mgc.nci.nih.gov/.
TITLE　National Institutes of Health, Mammalian Gene Collection (
JOURNAL　Unpublished (1999)
COMMENT　Contact: Robert Strausberg, Ph.D.
　　　Email: cgapbs-r@mail.nih.gov
　　　Tissue Procurement: The Cepko Laboratory
　　　cDNA Library Preparation: Life Technologies, Inc.
　　　cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
　　　DNA Sequencing by: Incyte Genomics, Inc.
　　　Clone distribution: MGC clone distribution information ca
　　　found through the I.M.A.G.E. Consortium/LLNL at:
　　　http://image.llnl.gov
　　　Plate: LLAM10435　row: h　column: 15
　　　High quality sequence stop: 468.
FEATURES　Location/Qualifiers
　source　1..650
　　　/organism="Mus musculus"
　　　/mol_type="mRNA"
　　　/db_xref="taxon:10090"
　　　/clone="IMAGE:4527038"
　　　/tissue_type="retina"
　　　/lab_host="DH10B (phage-resistant)"
　　　/clone_lib="NIH_MGC_94"
　　　/note="Organ: eye; Vector: PCMV-SPORT6; Site_1: 1
　　　Site_2: SalI; Cloned unidirectionally; oligo-dT 1
　　　Average insert size 3.3 kb. Library enriched for
　　　full-length clones and constructed by Life Techno
　　　Note: this is a NIH_MGC Library."

ORIGIN

Query Match　　　　41.9%;　Score 489.6;　DB 12;　Length 650;
Best Local Similarity　92.0%;　Pred. No. 1.6e-116;
Matches 550;　Conservative　0;　Mismatches　44;　Indels　4;　C

```
Qy 291 GAGCACAAGCAGCTGTGATGGACACAGTGAGTGGCTGGGAAGAGACCAAAATCAAC
       |||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db  25 GACCCCATTCAGGTGTGATGGGACACAGTGAGTGGCTGGGAAGAGACCAAAATCAA(

Qy 351 CCAGCCCTGCTGCCTACGACCCGCCAGATTGGGGAATTTACAGTCATCAGGGCTGG(
       ||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db  85 CCAGCCCTGCTGCCTACGACCGCCAGATTGGGGAATTTACAGTCATCAGGGCTGG(

Qy 411 ACTACCTGTACTGTCAGGTGCACTTTGATGAGGGAAAGGCTGTCTACCTGAAGCT(
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 145 ACTACCTGTACTGTCAGGTGCACTTTGATGAGGGAAAGGCTGTCTACCTGAAGCT(

Qy 471 TGCTGGTGAACGGTGCTGGCCCTGGCCCTGGCCCTGGGAAGAATTCTCAGCCACAGC
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 205 TGCTGGTGAACGGTGCTGGCCCTGGCCCTGGCCCTGGGAAGAATTCTCAGCCACAG(

Qy 531 GCTCTCCTGGGCCCCAGCTCCGTTGTGCCAGGTGTCTGGGCTGTTGCCGCTGCGG(
       |||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db 265 GCTCTCCTGGGCCCCAGCTCCGTTGTGTGCCAGGTGTCTGGGCTGTTGCCGCTGCGG(

Qy 591 GGTCTTCCCTTGGATCCGCACCCTCCCCTGGGCTCATCTTAAGGCTGCCCCCTT(
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 324 GGTCTTCCCTTGGATCCGCACCCTCCCCTGGGCTCATCTTAAGGCTGCCCCCTT(

Qy 651 CCTACTTTGGACTCTTTCAAGTTCACTGAGGGGCCTTGCTCTCCCAGATTCCTTA(
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 384 CCTACTTTGGACTCTTTCAAGTTCACTGAGGGGCCTTGCTCTCCCAGATTCCTTA(
```

AAGATATCAA; cell line R3327 5A, GGACTAGATC; cell
R3327 5P, CACGTGAGAT; duodenum, TGTGGTTCAT; prost
CCAGG.
TAG_TISSUE=duodenum
TAG_LIB=UI-R-EP0
TAG_SEQ=TGTGGTTCAT"

ORIGIN

Query Match   41.8%;   Score 488.8; DB 13;   Length 687;
Best Local Similarity   87.9%;   Pred. No. 2.6e-116;
Matches 590; Conservative   0; Mismatches 73;   Indels   8; G:

Qy   496 GGGCTGCCTGGAAGAATTCAGCCACAGCAGCAAGCTCCTGGGCCCCAGCTCC
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   687 GGGCTGCCTGGAAGAATTCAGCCACAGCAGCAAGCTCCTGGGCCCCAGCTCC

Qy   556 GTGCCAGGTGTCTGGGCTGTGCCCTGCGCCAGGGTCTTCCCTTCAAG
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   627 GTGCCAGGTGTCTGGGCTGTGCCTGCGCCAGGGTCTTCCCTTCGGACTCTTTCAAG

Qy   616 CCCCTGGGCTCATCTTAAGGCTGCCCCTTCCTAACCTACTTTGGACTCTTTCAAG
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   567 CCCCTGGGCTCATCTTAAGGCTGCCCCTTCCTTACCTACTTTGGACTCTTTCAAG

Qy   676 CTGAGGGGCCTTGCTCTCCCAGATTCCTAAACTTTCCC---TGGCTCCAGGAGC
           |||||||||||||||||||||||||||||||||||||| ||||||
Db   507 CTGAGGGGTCCTGCTCTCCCAGGTTCTTTAAACTTTAACCTAACAGCTCCCAGAGC

Qy   732 CCACACCTCCCTACCCCACCTCTCCACCCCCTCGCTGCTTCCTTGGTCCAG
           ||||||||||||||||||||||||||||||||||||||||||||
Db   447 CAGCACCTCCCCACCCCCAGCCCACTCCTCCACCCTNTAGCTGTTCCTTGCTCCAG

Qy   792 GT-CTCTCCTCAAAGGCAGCCAGAGCTTGTTCACATG-TTTCCATTCCACAGACGT
           || ||||||||||||||||||||||||||||||||| ||||||||||||||
Db   387 GTCTCTCTCCTTCCAAGGCAGCCAGCCAGGGCTTGATGACATGTTTTCCATTCCACAGACAT

Qy   850 TTGCTCTTCT-TAACATCCATCCCACCACAACTATCCACCTCACTAGCTCCCAAA
           | ||||||||| ||||||||||||||||||||||||||||||||||||||
Db   327 CCGTTCTTGTAACCTCCACCCCACCCACACCAGCTGCCCGCTCCAGAGCTCCCAAA

Qy   909 CTACTTATCCTGACTCCCCACCCACCCGACCACGTGTTTATTGACTTTGT
           |||||||||||||||||||||||||||||||||||||||
Db   267 CCACGTA-CCCGACTACCCCGCCCAACCCTGGACCCGTGTTTTTACTGACTATGT

Qy   969 CAGGCACTGAGATGGCTGACCTCGTGGCAGGAGCCAGAGAACCTGGGACTAGG
           ||||||||||||||||||||||||||||||||||||||||||
Db   208 CAGGCATCAAGAGATAGGCTGGACCTCGTGGCAGGAGCCAAAGAACCTGGGACTAGG

Qy   1029 AAGTTCCCAACTGTGAGGGGGAAGAGCTGGGGAACAAGCTCCTCCTGGATCCCTGT
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   148 AAGTTCCCAAATGTGAGGGAAAGAGATGGAGACAAGCTCCTCCCTGGATCCTTGT

Qy   1089 TTTGAAAAGATACATTTTATTATTTGTGACAAAAATGTTAAATGGATATTAAA
           ||||||||||||||||||||||||||||||||||||||||||||
Db   88 TTTGAAAAGATACTATTTTTATTATTATTTGTGACAAAAATGTTAAATGGATATTAAA

Qy   1149 ATAAATCATGA 1159
           |||||||||||
Db   28 ATAAATCATGA 18

RESULT 14
AW917574
LOCUS   AW917574   584 bp   mRNA   linear   EST 25
DEFINITION EST348878 Rat gene index, normalized rat, norvegicus, Bent
            Rattus norvegicus cDNA clone RGIEF49 5' end, mRNA sequence
ACCESSION   AW917574
VERSION   AW917574.1 GI:8083328
KEYWORDS   EST.
SOURCE   Rattus norvegicus (Norway rat)
  ORGANISM Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murin

---

GGCTCCAGGAGCATCACCACACCTTCCCTAGCCCACCCACTCCTCCACCCCCTC 770
           |||||||||| |||||||||||||| | ||||||||||||||||
GGGCTCCAGGAGCATCACCACCACCACCAGTACCCCACCACCAGTAACTCCACCCACTC 503

TCCTTGGTCCAGTCCTGT-CTCTCCTCAAAGGCAGCCAGAGCTTGTTCACATGT 829
           ||||||||||| |||||||||||||||||||||||||||||||
TCCTTGGTCCAGTCCTGTCCTCCTCATCAAAGGCAGACAGAGAGCTTGTACACATGTA 563

TCCACAGACGTATCCTTGCTCTTCTTAACATCCCATCCCACCACAACTATCCA 887
           | |||| ||| ||| ||| | ||||| ||||||||||||
--CCAAAGAGTATACTAGATCTACTTAAATAGCCATGCATGCAAAACAATATACACA 619

3   687 bp   mRNA   linear   EST 02-MAY-2002
0-coc-1-22-0-UI.s1 UI-R-EP0 Rattus norvegicus cDNA clone
0-coc-1-22-0-UI 3', mRNA sequence.
3
3.1 GI:20424898

norvegicus (Norway rat)
norvegicus
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

es 1 to 687)
,M.F., Lennon,G. and Soares,M.B.
zation and subtraction: two approaches to facilitate gene
ry
Res. 6 (9), 791-806 (1996)
7

. Soares, MB
ated Laboratory for Computational Genomics
ity of Iowa
ton Road, 4156 MEBRF, Iowa City, IA 52242, USA
9 335 8250
9 335 9565
bento-soares@uiowa.edu
quence contained an oligo-dT track that was present in the
cleotide that was used to prime the synthesis of first
cDNA and therefore this may represent a bonafide poly A
he sequence tag present in the cDNA between the NotI site
oligo-dT track served to identify it as a clone from the
zed duodenum library cDNA Library Preparation: M.B. Soares
ne distribution: clones will be available through Research
s (www.resgen.com)
mer: M13 Forward
es.

Location/Qualifiers
1..687
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-EP0-coc-1-22-0-UI"
/dev_stage="ADULT"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-EP0"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; UI-R-EP0 is a
subtracted cDNA library constructed according to Bonaldo,
Lennon and Soares, Genome Research, 6:791-806, 1996. First
strand cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into pT7T3-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dT)18 tail. The sequence tags for this
library are: distal colon, GAAGTGCTCC; osteoblast,

```
Db        542  AGCCAGNGCTTGATGACATGTTTTCCATTCCAGACACATATTC  584
```

RESULT 15

```
LOCUS        CB849011
DEFINITION   CB849011  MRA-0673 MOUSE ADULT RETINA Mus musculus cDNA 5', mRNA seq
ACCESSION    CB849011                        939 bp    mRNA    linear    EST 01
VERSION      CB849011.1  GI:34379595
KEYWORDS     EST.
SOURCE       Mus musculus (house mouse)
ORGANISM     Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1  (bases 1 to 939)
AUTHORS      Yu,J., Farjo,R., MacNee,S.P., Baehr,W., Stambolian,D.E. and
             Swaroop,A.
TITLE        Annotation and analysis of 10,000 expressed sequence tags
             developing mouse eye and adult retina
JOURNAL      Genome Biol. 4 (10), R65 (2003)
MEDLINE      22881944
PUBMED       14519200
COMMENT      Contact: Swaroop, A.
             Department of Ophthalmology and Visual Sciences
             Kellogg Eye Center, University of Michigan
             540 KEC, 1000 Wall St., Ann Arbor, MI 48105, USA
             Tel: 734 615 2246
             Fax: 734 647 0228
             Email: swaroop@umich.edu.
FEATURES             Location/Qualifiers
     source          1..939
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /tissue_type="retina"
                     /clone_lib="MOUSE ADULT RETINA"
                     /note="Vector: pSPORT1"
ORIGIN
```

```
Query Match             41.2%;    Score 480.8;  DB 14;   Length 939;
Best Local Similarity   90.4%;    Pred. No. 3.7e-114;
Matches  553;  Conservative    0;  Mismatches  52;   Indels    7;   Gaps
```

```
Qy   225  GGAAGGGCGCGGCCTCGCCCAGCTATTGCAGCCCATTATGAGTTCATCCTCGGCCA
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    25  GGAAAGTGTGTCCGCTGCAGGTACCGTCCGGAATTCCCGGGTCGACCCACCGCGA

Qy   285  AGGATGGAGCACAAGCAGGTGTGATGGAGCAGGTGAGTGGCTGGGAAGAGACCAAA
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    85  AGGATGGAGCACAAGCAGGTGTGATGGAGCAGGTGAGTGGCTGGGAAGAGACCAAA

Qy   345  ACAGCTCCAGCCCTCGCGCTACGACCGCCAGATTGGGGAATTACAGTCATCAGC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   145  ACAGCTCCAGCCCTCGCGCTACGACCGCCAGATTGGGGAATTTACAGTCATCAGC

Qy   405  GGCTCTACTACCTGTACTGTCAGGTGCACTTTGGAGGAAAGGCTGTCTACCTC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   205  GGCTCTACTACCTGTACTGTCAGGTGCACTTTGGAGGAAAGGCTGTCTACCTC

Qy   465  TGGACTTGCTGGTGAACGGTGTGCTGGCCCTGAGGGGCTTGCTC
          ||||||||||||||||||||||||||||||||||||||||||||
Db   265  TGGACTTGCTGGTGAACGGTGTGCTGGCCCTGAGGGGGTCCTGCTC

Qy   525  CAGCAAGCTCTCCTGGGCCCCAGCTCCGTTTGTGCCAGGTGTCTGGGCTGTTGCC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   325  CAGCAAGCTCTCCTGGGCCCCAGCTCCGTTTGTGCCAGGTGTCTGGGCTGTTGCC

Qy   585  GGCCAGGGTCTTCCTTCGGATCCGGCACCCTCCCTCCGCCTCATCTTAAGGCTGCC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   385  GGCCAGGGTCTTCCTTCGGATCCGGCACCCTCCCTCCGCCTCATCTTAAGGCTGCC

Qy   645  TCCTAACCTACTTTGGACTCTTTCAAGTTCACTGAGGGGCCTTGCTCTCCCAGATT
          |||||||||||||||||||||||||||||||||||||||||| ||||
```

```
ses  1 to 584)
H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,
age,A.R.and Adams,M.D.
nome Project: Generation of a Rat EST (REST) Catalog & Rat
ndex
ished (1998)
t: Lee, NH
stitute for Genomic Research
Medical Center Drive, Rockville, MD 20850, USA
301)-838-3529
301)-838-0208
nhlee@tigr.org
lone is available through the ATCC, contact the ATCC
3-365-2700 for further information
imer: M13 Reverse
            Location/Qualifiers
     1..584
            /organism="Rattus norvegicus"
            /mol_type="mRNA"
            /db_xref="taxon:10116"
            /clone="RGIEF49"
            /tissue_type="mix - brain, ovary, placenta, kidney, lung,
            liver, embryo, heart, muscle, spleen"
            /lab_host="SOLR"
            /clone_lib="Rat gene index, normalized rat, norvegicus,
            Bento_Soares"
            /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
            XhoI; Estimated insert size approx.1 kb"
```

```
arity        41.2%;   Score 481.2;   DB 10;   Length 584;
             91.3%;   Pred. No. 2.3e-114;
onservative    0;   Mismatches  46;   Indels    5;   Gaps    2;
```

```
Qy        GGCCAGGACAGGATGGGAGCACAAGCAGCAGGTGTGATGGGACAGTGGCTGGGAA  331
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        AGCCAGGACAGGATGGGAGCACAAGCAGGCAGGTGTGATGGGACAGTGGCTGGGAA  61

Qy        CCAAAATCAACAGCTCCAGCCCTCGGCTACGACCGCCAGATTGGGGAATTTACA  391
          ||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db        CCAAAATCAACAGCTCCAGCCCCTCGCGCTATGACCGGCATCAGCGGGAATTTACG  121

Qy        TCAGGGGTGGGCTTCTACTACCTGTACTGTCAGGTCACTTTGATGAGGAAAGGCT  451
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        TCAGGGGTGGGCTTCTACTACCTGTACTGTCAGGTGCACTTTGATGAGGGGAAGGCA  181

Qy        ACCTGAACTGAACTTGCTGGTGAACGGTGTCTGGCCCTGCGCCTGGGAAGAA  511
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        ACCTGAACTGAACTTGCTGGTGAATGGTGTCGGTGCGCCTGCGCCTGCTGGGAAGAA  241

Qy        CAGCCACAGCAGCAGCAAGCTCTCTCGGGCCCCAGCTCCGTTTGTGCCAGGTGTCTGGG  571
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        CAGCCACAGCAGCAGCAGCTCTCCAGCCCCTCGGGCCCCAGCTCCGTCGTGTGCCAGGTGTCTGGG  301

Qy        TGCCGCTGCGGCCGGAGGGTCTTCCCTTCGGATCCGCACCCTCCCTCCGGGTCATCTT  631
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        TGCCTCTGCGGGCCGGAGGTCTTCCCTTCGGATCCGTACCCTCCCTCCCGGGCTCATCTT  361

Qy        CTGCCCCCTTCCTAACCTACTTTGGACTCTTTCAAGTTCACTGAGGGGCCTTGCTC  691
          ||||||||||||||||||||||||||||||||||||||||||| |||||
Db        CTGCCCCCTTCCTTACCTACCTGGACTCTNTCAAGTTCACTGAGGGGTCCTGCTC  421

Qy        AGATTCCTAAACTTTCCC----TGGCTCCAGGAGCATCACCACACCTCCCTACCC  747
          ||||||||||  |||||||    |||||||||||||  ||||  ||||||||||||
Db        AGGTTCTTTAAACNTAACCTAACGCCCCAGAGCACCAGCACCTCCCCACCC  481

Qy        CCCACTCCTCCACCCCCTCGCTCAGTCCTTGGTCCAGTCCTGTCTCCTCAAAGGC  807
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        CCCACTCCTCCACCCCCACCCTTAGCTGTTTCCTTGCTCCAGTCCTGTCTCCTCTTAAGGC  541

Qy        AGAGCTTGTTCACATG-TTTCCATTCCAGACACGTATCC  849
          |||||| ||||||| || |||| |||||
```

```
ACCTACTTTGGACTCTTCAAGTTCACTGAGGGCCTTGCTCTCCCAGATACCTT 504

TTCCCTGGCTCC-AGGAGCATCACCACCACCTCCCTACCCCACCCCCACTCCTCCA 763
|||||||||||| ||||||||||||||||||| |||||||||| |||||||| |||
TTNCCTGCTCCAAGGAGCATCACCACCATCACCCCACCCCCACTNCTGCA 564

T-CGCTGCTCCTTGGTCCAGTCC---TGTCTCCTCAAAGGCA--GCCAGAGCT 817
| ||||||||||| ||||||||  || ||||||||| |  |||||||
TNNNCTGCTCCTTGGGACCAANTCCTGTTCCTCNCTNCTNAAAGGCANNNCAANAGCT 624

ACATGTT 829
|||||||
ACATGTT 636
```

pril 7, 2004, 23:15:14
secs

GenCore version 5.1.6
opyright (c) 1993 - 2004 Compugen Ltd.

c search, using sw model

il 7, 2004, 17:30:19 ; Search time 446.331 Seconds
(without alignments)
1117.066 Million cell updates/sec

09-245-198A-1

gtgctgagcctggcctgg........ataaatcatgattctcctc 1168

NTITY_NUC
op 10.0 , Gapext 1.0

3863 seqs, 2124099041 residues

s satisfying chosen parameters: 6747726

th: 0
th: 2000000000

nimum Match 0%
ximum Match 100%
sting first 45 summaries

Geneseq_29Jan04:*
geneseqn1980s:*
geneseqn1990s:*
geneseqn2000s:*
geneseqn2001as:*
geneseqn2001bs:*
geneseqn2002s:*
geneseqn2002as:*
geneseqn2003as:*
geneseqn2003bs:*
geneseqn2003cs:*
geneseqn2004s:*

the number of results predicted by chance to have a
than or equal to the score of the result being printed,
d by analysis of the total score distribution.

SUMMARIES

ry ch	Length	DB	ID	Description
.0	1168	2	AAV18599	Aav18599 Mus muscu
.0	1239	9	ADC97713	Adc97713 Murine FL
.9	701	3	AAX23425	Aax23425 Mouse TNR
.3	1353	3	AAA49717	Aaa49717 Human PRO
.8	1353	6	ABK40255	Abk40255 cDNA enco
.8	1421	4	AAX56000	Aax56000 Human tum
.4	1306	7	ACC57587	Acc57587 Polynucle
.4	1306	7	ACC57901	Acc57901 Human TWE
.9	1306	6	ADC35205	Adc35205 Human cDN
.9	1364	6	ABK34881	Abk34881 Human CDN
.6	1373	2	AAV18600	Aav18600 Homo sapi
.2	1236	4	AAV47613	Aav47613 TNF relat
.2	1236	4	AAD04350	Aad04350 Human TRE
.8	1030	2	AAX23424	Aax23424 Human TNR
.7	898	4	AAS03964	Aas03964 Expressio
.5	412	9	ADB56326	Adb56326 Toxicity-
.2	408	7	ABX37032	Abx37032 Bovine ES
.1	282	6	AAT22190	Aat22190 Human gen
.0	195	8	ABK29540	Abk29540 Colon ade
.6	493	8	ACH34013	Ach34013 Human end
.6	65	6	ABN55975	Abn55975 Mouse spl
.2	1064	6	ABT09678	Abt09678 Human PAL
.9	264	7	ABX52254	Abx52254 Bovine ES

c	24	52.6	4.5	2000	7	ADA71938	Ada71938 Ric
c	25	46.8	4.0	1117	9	ADC86688	Adc86688 Hur
c	26	46.6	4.0	3133	6	ADC86738	Adc86738 Hur
	27	46.6	4.0	53522	6	AAD30228	Aad30228 Hur
	28	46.6	4.0	53526	2	AAT94101	Aat94101 Hur
	29	46.6	4.0	53577	2	AAT18551	Aat18551 Hur
	30	46.6	4.0	53577	2	AAT94108	Aat94108 Hur
	31	45.8	3.9	1065	6	ABT09682	Abt09682 Hur
	32	45.8	3.9	105325	9	ABK94407	Abk94407 DN
c	33	45.6	3.9	5452	9	ADC86736	Adc86736 Hur
	34	45	3.9	1337	2	AAZ17263	Aaz17263 Hur
c	35	44.8	3.8	1000	2	AAA02484	Aaa02484 Hur
c	36	44.8	3.8	3163	9	ADC87060	Adc87060 Hur
c	37	44.8	3.8	12733	6	ABK98631	Abk98631 Vec
c	38	44.8	3.8	12733	6	ACD13882	Acd13882 L.
c	39	44.8	3.8	12739	6	ABK98592	Abk98592 Vec
c	40	44.8	3.8	12739	8	ACD13843	Acd13843 Pla
c	41	43.6	3.7	16167	8	ABL33083	Abl33083 Hur
c	42	43.6	3.7	16167	6	ABL34529	Abl34529 Hur
c	43	43.6	3.7	16167	6	ABL70254	Abl70254 Che
	44	43.2	3.7	324	6	ABL51826	Abl51826 Se
	45	43.2	3.7	720	4	AAH05001	Aah05001 Hur

ALIGNMENTS

RESULT 1
AAV18599
ID   AAV18599 standard; cDNA; 1168 BP.
XX
AC   AAV18599;
XX
DT   21-JUL-1998 (first entry)
XX
DE   Mus musculus tumour necrosis factor related ligand (TRELL) gene.
XX
KW   TRELL; tumour necrosis factor related ligand; tnf; treatment; can
KW   autoimmune disease; immune system; stimulation; suppression;
KW   graft rejection; ds.
XX
OS   Mus musculus.
XX
FH   Key   Location/Qualifiers
FT   CDS   2..679
FT   /*tag=   a
FT   /note= "tumour necrosis factor related ligand"
XX
PN   WO9805783-A1.
XX
PD   12-FEB-1998.
XX
PF   07-AUG-1997; 97NO-US013945.
XX
PR   07-AUG-1996; 96US-0023541P.
PR   18-OCT-1996; 96US-0028515P.
PR   18-MAR-1997; 97US-0040820P.
XX
PA   (BIOJ) BIOGEN INC.
PA   (UYGE-) UNIV GENEVA FACULTY MEDICINE.
XX
PI   Chicheportiche Y, Browning JL;
XX
DR   WPI; 1998-145619/13.
DR   P-PSDB; AAW47524.
XX
PT   Tumour necrosis factor related ligand - useful for, e.g. treating
PT   auto-immune disease and immune responses to tissue grafts.
XX
PS   Claim 2; Page 45-46; 69pp; English.
XX
CC   The sequence is that encoding mouse tumour necrosis factor related
CC   (TRELL). TRELL or active fragments can be included with a carrier

l compositions to treat cancer, autoimmune diseases or
uses to tissue grafts, or to stimulate or suppress the immune
useful to screen for TRELL receptors, by labelling with a
bel and screening compositions for binding. Agents
ith TRELL-receptor binding can also be screened for, can
istered, optionally with interferon-gamma, to induce cell
t, suppress or alter immune responses (especially involving
rcinoma cells) involving a signal pathway between TRELL and
The DNA sequence can be used in gene therapy for TRELL-
ders in mammals (especially humans), e.g. tumours,
d inflammatory diseases or inherited genetic disorders, by
nto cells, and expressing therapeutically effective amounts
e.g. a virus comprising a gene encoding TRELL. It may also
the preparation of prepare probes for screening
etic DNAs for TRELL-encoding sequences and for antisense

BP; 242 A; 360 C; 298 G; 268 T; 0 U; 0 Other;

arity 100.0%; Score 1168; DB 2; Length 1168;
100.0%; Pred. No. 8.6e-306;
onservative 0; Mismatches 0; Indels 0; Gaps 0;

```
                                 60
Qy   CTGAGCCTGGCCTGGGCGCTGGCCTGCCTTGGCCTCCTGCTGGTCGTGTGTCAGCCT
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   CTGAGCCTGGCCTGGGCGCTGGCCTGCCTTGGCCTCCTGCTGGTCGTGTGTCAGCCT   60

                                 120
Qy   AGCTGGGCAACGCTGTCTGCCCAGGAGCCTTCTCAGGAGGAGCTGACAGCAGAGGA
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   AGCTGGGCAACGCTGTCTGCCCAGGAGCCTTCTCAGGAGGAGCTGACAGCAGAGGA   120

                                 180
Qy   CGGGAGCCCCCTGAACTGAATCCCCAGACAGAGAGGAAAGCCAGGATGTGGTACCTTT
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   CGGGAGCCCCCTGAACTGAATCCCCAGACAGAGAGGAAAGCCAGGATGTGGTACCTTT   180

                                 240
Qy   GAACAACTAGTCCGGCCTCGAAGAGTGTCCTAAAGGCCGGAAGGCGCGGCCTCG
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   GAACAACTAGTCCGGCCTCGAAGAGTGTCCTAAAGGCCGGAAGGCGCGGCCTCG   240

                                 300
Qy   GCTATTGCAGCCATTATGAGGTTCATCCTCGGCCAGACACAGAGTGAGCACAAGC
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   GCTATTGCAGCCCATTATGAGGTTCATCCTCGGCCAGACACAGAGTGAGCACAAGC   300

                                 360
Qy   GTGGATGGGACAGTGAGTGGCTGGGAAGAGACCAAAATCAACAGCTCCAGCCCTCT
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   GTGGATGGGACAGTGAGTGGCTGGGAAGAGACCAAAATCAACAGCTCCAGCCCTCT   360

                                 420
Qy   TACGACCGGCAGATTGGGAATTTACAGTCATCAGGGCTGGGCTCTACTACCTGTA
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   TACGACCGGCAGATTGGGAATTTACAGTCATCAGGGCTGGGCTCTACTACCTGTA   420

                                 480
Qy   CAGGTGCACTTTGATGAGGGAAAAGCTGTCTACCTGAAGCTGGACTTGCTGGTGAA
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   CAGGTGCACTTTGATGAGGGAAAAGCTGTCTACCTGAAGCTGGACTTGCTGGTGAA   480

                                 540
Qy   GTGCTGCGCCCTGCCCTGCGCTGGAGAATTCTCAGCCACAGCAGCAAGCTCTCCTGG
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   GTGCTGCGCCCTGCCCTGCGCTGGAGAATTCTCAGCCACAGCAGCAAGCTCTCCTGG   540

                                 600
Qy   CAGCTCCGTTTGTGCCAGGTGTCTGGGCTGTTGCGCCTGCGGCCAGGGTCTTCCCT
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   CAGCTCCGTTTGTGCCAGGTGTCTGGGCTGTTGCGCCTGCGGCCAGGGTCTTCCCT   600

                                 660
Qy   GATCCGCACCCTCCCTGGCTCATCTTAAGGCTGCCCCCTTCTAACCTACTTTGG
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   GATCCGCACCCTCCCTGGCTCATCTTAAGGCTGCCCCCTTCTAACCTACTTTGG   660

                                 720
Qy   CTTTCAAGTCACTGAGGGGCCTTGCTTCTCCCAGATTCCTTAAACTTTCCCTGGCTC
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   CTTTCAAGTTCACTGAGGGGCCTTGCTTCTCCCAGATTCCTTAAACTTTCCCTGGCTC   720

                                 780
Qy   GAGCATCACCACACCTCCCTACCCCCACCCCACTCCTCCACCCCCTGCTCCTTT
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
                                 780
Db   CAGGAGCATCACCACACCTCCCTACCCCCACCCCACTCCTCCACCCCCTCGCTGC

                                 840
Qy   GGTCCAGTCCTGTCTCCTCCTCAAAGGCCCAGAGCTTGTTCACATGTTTCCATT(
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   GGTCCAGTCCTGTCTCCTCCTCAAAGGCCCAGAGCTTGTTCACATGTTTCCATT(  781

                                 900
Qy   GACGTATCCTTGCTCTTCTTAACATCCCATCCACCAACTATCCACCTCACTAC
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   GACGTATCCTTGCTCTTCTTAACATCCCATCCACCAACTATCCACCTCACTAC  841

                                 960
Qy   CAAAGCCCCTATTATCCCTGACTCCCCACCCACTCACCCGACCACGTGTTTAT
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   CAAAGCCCCTACTTATCCCTGACTCCCCACCCACTCACCCGACCACGTGTTTAT  901

                                1020
Qy   TTGTGCACCAGGCACTGAGATGGGCTGAGGGGGGAAGAGCTGGGAGGAGCCAAGACCTCCTCCTG(
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   TTGTGCACCAGGCACTGAGATGGGCTGACCTGGTGGGGGCAGGAGGAGCCAAGACCTCCTCCTG(  961

                                1080
Qy   TAGGCCAGAAGTTCCCAACTGTGAGGGGAAGAGCTGGGACAAGAGCTCCTCCTG(
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   TAGGCCAGAAGTTCCCAACTGTGAGGGGAAGAGCTGGGACAAGAGCTCCTCCTG(  1021

                                1140
Qy   CTGTGGATTTTGAAAAGATACATATTTTTATTATTGTGACAAAATGTTAAATG(
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   CTGTGGATTTTGAAAAGATACATATTTTTATTATTGTGACAAAATGTTAAATG(  1081

                                1168
Qy   TAAAGAGAATAAATCATGATTTCTCTTC  1168
     ||||||||||||||||||||||||||||
Db   TAAAGAGAATAAATCATGATTTCTCTTC  1168
```

RESULT 2
ADC97713
ID   ADC97713 standard; cDNA; 1239 BP.
XX
AC   ADC97713;
XX
DT   15-JAN-2004 (first entry)
XX
DE   Murine FL-TWEAK coding sequence.
XX
KW   Murine; FL-TWEAK; TNF relatedness and weak ability to induce cel
KW   TNF; Tumour Necrosis Factor; TWEAK; fibrosis; cardiac disease;
KW   liver disease; lung disease; kidney disease; skin disease;
KW   skeletal muscle disease; adipose tissue disease;
KW   gastrointestinal tract disease; pancreatic disease;
KW   reproductive organ disease; neural disease; cartilage disease;
KW   bone disease; connective tissue disease; cellular death; hepatot
KW   dermatological; gastrointestinal; osteopathic; gene; ss.
XX
OS   Mus sp.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..750
FT   /*tag= a
FT   /product= "FL-TWEAK"
XX
PN   WO200306311-A2.
XX
PD   23-OCT-2003.
XX
PF   09-APR-2003; 2003WO-US011350.
XX
PR   09-APR-2002; 2002US-0371611P.
XX
PA   (BIOJ) BIOGEN INC.
XX
PI   Burkly L, Jakubowski A, Zheng T, Hahm K;
XX
DR   WPI; 2003-845256/78.
DR   P-PSDB; ADC97712.
XX
PT   Treating a TWEAK-related condition, e.g. liver, gastrointestinal

ic, cartilage or neural tissue condition in a subject
ministering to the subject a TWEAK agonist or antagonist.

) ID NO 2; 120pp; English.

equence is the coding sequence for murine transmembrane FL-
atedness and weak ability to induce cell death, where TNF
osis Factor). TWEAK is a member of the TNF family. TWEAK
tagonists are useful for treating a TWEAK-related
, fibrosis; cardiac disease; liver disease; lung disease;
; skin disease; skeletal muscle disease; adipose tissue
ointestinal tract disease; pancreatic disease; reproductive
neural disease; cartilage disease; bone disease;
ssue disease; cellular death; and a pathological condition
xpressing a TWEAK receptor.

BP; 249 A; 386 C; 331 G; 273 T; 0 U; 0 Other;

rity   100.0%;   Score 1168;   DB 9;   Length 1239;
         100.0%;   Pred. No. 8.8e-306;
onservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
                CTGAGCCTGGGCCTGGCGCTGGCCTGCCTTCCTGCTTCGTGGTCAGCCT    60
                ||||||||||||||||||||||||||||||||||||||||||||||||||
                CTGAGCCTGGGCCTGGCGCTGGCCTGCCTTCCTGCTTCGTGGTCAGCCT    131

                AGCTGGGCAACGCTGTCTGCCCAGGAGCCTTCTCAGGAGGAGCTGACAGCAGAGGA    120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                AGCTGGGCAACGCTGTCTGCCCAGGAGCCTTCTCAGGAGGAGCTGACAGCAGAGGA    191

                CGGGAGCCCCCTGAACTGAATCCCCAGACACAGAGGAAAGCCAGGATGTGGTACCTTT    180
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                CGGGAGCCCCCTGAACTGAATCCCCAGACACAGAGGAAAGCCAGGATGTGGTACCTTT    251

                GAACAACTAGTCCGGCCTCGAAGAAGTGCTCCTAAAGGCCGGAAGGCGCGGCCTCG    240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                GAACAACTAGTCCGGCCTCGAAGAAGTGCTCCTAAAGGCCGGAAGGCGCGGCCTCG    311

                GCTATTGCAGCCATTATGAGGTTCATCCTCGGCCAGGACAGGATGGAGCACAAGC    300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                GCTATTGCAGCCATTATGAGGTTCATCCTCGGCCAGGACAGGATGGAGCACAAGC    371

                GTGGATGGGACAGTGAGTGGCTGGGAAGAGACCAAAATCAACAGCTCCAGCCCTCT    360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                GTGGATGGGACAGTGAGTGGCTGGGAAGAGACCAAAATCAACAGCTCCAGCCCTCT    431

                TACGACCGGCCAGATTGGGGAATTACAGTCATCAGGGCTGGGCTCTTACTACCTGTA    420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                TACGACCGGCCAGATTGGGGAATTACAGTCATCAGGGCTGGGCTCTTACTACCTGTA    491

                CAGGTGCACTTTGATGAGGGAAAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGAA    480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                CAGGTGCACTTTGATGAGGGAAAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGAA    551

                GTGCTTGGCCCTGCGCGCTGCCTGCCTGAAGAATTCTCAGCCACAGCAGCAAGCTCCCTGG    540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                GTGCTTGGCCCTGCGCGCTGCCTGCCTGAAGAATTCTCAGCCACAGCAGCAAGCTCCCTGG    611

                CAGCTCCGTTTGTGCCAGGTGTCTGGGCTGTTGCCGCTGCGGCCAGGGTCTTCCCT    600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                CAGCTCCGTTTGTGCCAGGTGTCTGGGCTGTTGCCGCTGCGGCCAGGGTCTTCCCT    671

                ATCCGCACCCCTCCCCCTGGGCTCATCTTAAGGCTGCCCCCCTTCCTAACCTACTTTGG    660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                ATCCGCACCCCTCCCCCTGGGCTCATCTTAAGGCTGCCCCCCTTCCTAACCTACTTTGG    731

                TTTCAAGTCACTGAGGGGCCTTGCTCTCCCAGATTCCTTAAACTTTCCCTGGCTC    720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                TTTCAAGTCACTGAGGGGCCTTGCTCTCCCAGATTCCTTAAACTTTCCCTGGCTC    791

                AGCATCACCACACCTCCCTACCCCACCCCCACTCCTCCACCCCCTCGCTGCTCCTT    780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                AGCATCACCACACCTCCCTACCCCACCCCCACTCCTCCACCCCCTCGCTGCTCCTT    851
```

Qy   781   GGTCCAGTCCTGTCTCTCCTCAAAGGCAGCAGCCAGAGCTTGTTCACATGTTCCATTC
Db   852   GGTCCAGTCCTGTCTCTCCTCAAAGGCAGCAGCCAGAGCTTGTTCACATGTTCCATTC

Qy   841   GACGTATCCTTGCTCTTCTTAACATCCATCCCACCACACTATCCACCTCACTAG
Db   912   GACGTATCCTTGCTCTTCTTAACATCCATCCCACCACACTATCCACCTCACTAG

Qy   901   CAAAGCCCTACTTATCCTGACTCCCCACCACTCACCCGACCACGTGTTTATT
Db   972   CAAAGCCCTACTTATCCTGACTCCCCACCACTCACCCGACCACGTGTTTATT

Qy   961   TTGTGCACCAGGCACTGAGATGGGCTGGACCTGGTGGCAGGAAGCCAGAGAACCTG
Db   1032   TTGTGCACCAGGCACTGAGATGGGCTGGACCTGGTGGCAGGAAGCCAGAGAACCTG

Qy   1021   TAGGCCAGAAGTTCCCAACTGTGAGGGGGAAGAGCTGGGGACAAGCTCCTCCTGG
Db   1092   TAGGCCAGAAGTTCCCAACTGTGAGGGGGAAGAGCTGGGGACAAGCTCCTCCTGG

Qy   1081   CTGTGGATTTTGAAAAGATACTATTTTATTATTGTGACAAAATGTTAAATGG
Db   1152   CTGTGGATTTTGAAAAGATACTATTTTATTATTGTGACAAAATGTTAAATGG

Qy   1141   TAAAGAGAATAAATCATGATTTCTCTTC   1168
Db   1212   TAAAGAGAATAAATCATGATTTCTCTTC   1239

RESULT 3
AAX23425
ID   AAX23425 standard; DNA; 701 BP.
XX
AC   AAX23425;
XX
DT   18-JUN-1999   (first entry)
XX
DE   Mouse TNRL3 DNA.
XX
KW   Tumour necrosis factor receptor; signal transducer molecule; TNF;
KW   developmental abnormality; gestational abnormality; prostate ca
KW   APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy;
KW   cytoplasmic domain; immunogen; antibody preparation; breast carci
KW   apoptosis; mouse; ss.
XX
OS   Mus sp.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..636
FT   /*tag=   a
FT   /product= "TNRL3"
XX
PN   WO9911791-A2.
XX
PD   11-MAR-1999.
XX
PF   04-SEP-1998;   98WO-US018393.
XX
PR   05-SEP-1997;   97US-00924634.
XX
PA   (UNIW) UNIV WASHINGTON.
XX
PI   Chaudhary PM;
XX
DR   WPI; 1999-205191/17.
DR   P-PSDB; AAW93591.
XX
PT   New Tumor Necrosis Factor family receptor polypeptides and ligand
PT   useful for diagnosis and treatment of prostate cancer and develop
PT   or gestational abnormalities.
XX
PS   Example VII; Fig 13B; 156pp; English.

n describes isolated Tumor Necrosis Factor (TNF) family
peptides: APO4, APO6, APO8 and APO9 or their active
d isolated TNF related ligands 1 and 3 (TNRl1 and TNRl3) or
fragments. APO4 is useful for diagnosing prostate cancer by
evels of APO4 in an individual. Prostate cancer can also be
APO4 selective binding agents linked to a therapeutic
polypeptides are also useful for identifying selective
s, useful in diagnosis/treatment of disease by binding of
polypeptide/active fragment which is extracellular, or
the cell surface. The binding is preferably performed in
polypeptides/ active fragments are also useful for screening
and antagonists by binding and observing the changer in APO4
disease are also identified using APO4 polypeptides/active
ective pharmacological agents useful in diagnosis or
l APO4 signal transducer molecules that specifically interact
asmic domain of APO4 and detecting a change in level of APO4
a method is performed in vivo or in vitro. APO polypeptides
l as immunogens for preparing antibodies. APO4 is also
agnosis/treatment of developmental or gestational
. APO8 was transfected to human breast carcinoma cell line
duced apoptosis

BP; 139 A; 210 C; 203 G; 149 T; 0 U; 0 Other;

arity   59.9%;   Score 699.4;   DB 2;   Length 701;
onservative   99.9%;   Pred. No. 5.8e-179;
            0; Mismatches   1; Indels   0; Gaps   0;

Qy   TCGTGGTCAGCCTGGGGAGCTGGGCAACGCTGTCTGCCCAGGAGCCTTCTCAGGAG 103
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   TCGTGGTCAGCCTGGGGAGCTGGGCAACGCTGTCTGCCCAGGAGCCTTCTCAGGAG 60

Qy   TGACAGCAGAGGACCGCCGGGAGCCCCCTGAACTGAATCCCAGACACAGAGGAAAGC 163
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   TGACAGCAGAGGACCGCCGGGAGCCCCCTGAACTGAATCCCAGACACAGAGGAAAGC 120

Qy   ATGTGGTACCTTTCTTGGAACAACTAGTCCGGCCTCGAAGAAGTGCTCCTAAAGGC 223
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   ATGTGGTACCTTTCTTGGAACAACTAGTCCGGCCTCGAAGAAGTGCTCCTAAAGGC 180

Qy   AGGGCGGCCTCGCCGAGCTATTGCAGCCCATTATGAGGTTCATCCTCGGCCAGGA 283
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   AGGGCGGCCTCGCCGAGCTATTGCAGCCCATTATGAGGTTCATCCTCGGCCAGGA 240

Qy   ATGAGCACAAGCAGGTGTGGATGGGACAGTGAGTGAGTGGCTGGGAAGAGACCAAAATC 343
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   ATGGAGCACAAGCAGGTGTGGATGGGACAGTGAGTGAGTGGCTGGGAAGAGACCAAAATC 300

Qy   AGCTTCCAGCCCTGCGTCTACGACCGCCAGATTGGGGAATTACAGTCATCAGGGCT 403
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   AGCTTCCAGCCCTGCGTCTACGACCGCCAGATTGGGGAATTACAGTCATCAGGGCT 360

Qy   TCTACTACCTGTACTGTCAGGTGCACTTTGATGAGGGAAAGGCTGTCTACCTGAAG 463
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   TCTACTACCTGTACTGTCAGGTGCACTTTGATGAGGGAAAGGCTGTCTACCTGAAG 420

Qy   ACTTGCTGGTGAACGGTGTGCTGGCCCTGCGCTGCTGCTGGAAGAATTCTCAGCCACA 523
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   ACTTGCTGGTGAACGGTGTGCTGGCCCTGCGCTGCTGCTGGAAGAATTCTCAGCCACA 480

Qy   CAAGCTCTCTGGGCCCCAGCTCCGTTGTGCCAGGTGTCTGGGCTGTTGCCGCTG 583
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   CAAGCTCTCTGGGCCCCAGCTCCGTTGTGCCAGGTGTCTGGGCTGTTGCCGCTG 540

Qy   CAGGGTCTTCCTTCGGATCCGACACCTCCCTGGGCTCATCTTAAGGCTGCCCCCC 643
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   CAGGGTCTTCCTTCGGATCCGACACCTCCCTGGGCTCATCTTAAGGCTGCCCCCC 600

Qy   CTAACCTACTTTGGACTCTTTCAAGTTCACTGAGGGGCCTTGCTCTCCCAGATTCCT 703
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   CTAACCTACTTTGGACTCTTTCAAGTTCACTGAGGGGCCTTGCTCTCCCAGATTCCT 660

Qy   704 TAAACTTTCCCTGGCTCCAGGAGCATCACCACACCTCCCTA 744
           |||||||||||||||||||||||||||||||||||||||||
Db   661 TAAACTTTCCCTGGCTCCAGGAGCATCACCACACCTCCCTA 701

RESULT 4
AAA49717
ID   AAA49717 standard; cDNA; 1353 BP.
XX
AC   AAA49717;
XX
DT   25-SEP-2000   (first entry)
XX
DE   Human PRO207 cDNA clone DNA30879-1152.
XX
KW   PRO207; human; antitumour; tumour; therapy; cytostatic; breast ca
KW   ovarian cancer; renal cancer; colorectal cancer; uterine cancer;
KW   prostate cancer; lung cancer; bladder cancer;
KW   central nervous system cancer; melanoma; leukaemia; neoplasm; ss
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   58..807
FT   /*tag= a
FT   sig_peptide   58..177
FT   /*tag= b
FT   mat_peptide   178..804
FT   /*tag= c
XX
PN   WO200037638-A2.
XX
PD   29-JUN-2000.
XX
PF   02-DEC-1999;   99WO-US028565.
XX
PR   22-DEC-1998;   98US-0113296P.
PR   08-MAR-1999;   99WO-US005028.
PR   21-APR-1999;   99US-0130232P.
PR   28-APR-1999;   99US-0131445P.
PR   14-MAY-1999;   99US-0134287P.
PR   20-JUL-1999;   99US-0144758P.
PR   26-JUL-1999;   99US-0145698P.
PR   15-SEP-1999;   99WO-US021090.
PR   15-SEP-1999;   99WO-US021547.
XX
PA   (GETH) GENENTECH INC.
XX
PI   Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA
PI   Napier MA, Pitti RM, Wood WI;
XX
DR   WPI; 2000-442668/38.
DR   P-PSDB; AAY95338.
XX
PT   Novel composition to inhibit neoplastic cell growth or for treat:
PT   in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,
PT   PRO224, PRO328, PRO301, PRO526, PRO328, PRO356, PRO509 or PRO866
XX
PS   Claim 20; Fig 3; 172pp; English.
XX
CC   The present sequence is that of cDNA clone DNA30879-1152 (ATCC 2(
CC   encoding human PRO207 (see AAY95338), which shows homology to se'
CC   members of the tumour necrosis factor family, especially human
CC   lymphotoxin (23.4%). The cDNA was identified in a foetal kidney (
CC   library following Apo-2 ligand. A claimed method for inhibiting '
CC   homology to human Apo-2 ligand. A claimed method for inhibiting '
CC   growth of a tumour cell comprises exposing the tumor cell to PRO:
CC   PRO207, PRO320, PRO219, PRO221, PRO224, PRO328, PRO301, PRO526,
CC   PRO356, PRO509 or PRO866 (see AAY95337-49), their agonists or ch
CC   polypeptides incorporating them. The tumour is especially a canc
CC   selected from breast, ovarian, renal, colorectal, uterine, prost.
CC   lung, bladder and central nervous system cancer, melanoma and le
CC   Nucleic acids encoding PRO179 etc. are used in the recombinant p:

our polypeptides

BP; 257 A; 443 C; 389 G; 264 T; 0 U; 0 Other;

53.8%;   Score 628.6; DB 3; Length 1353;
rity   76.2%;   Pred. No. 1.2e-159;
nservative   0; Mismatches 219; Indels   73; Gaps   10;

```
GAGCCTGGGCCTGGCCTGCGCTGCCCTGGCCCTGCTGGTCGTGGTCAGCCTG    61
|||||||||||||||||||||||||||||||||||||||||||||||||||||
GGGCCTGGGCCTGGCCTGCGCTGCCCTGCCCTGCTGGCCGTGGTCAGTTTG    189

CTGGGCAACGCTGTCTGCGCCAGGAGCCTTCTCAGGAGGAGCTGACAGCAGAGGAC    121
|| |||| ||||||||||||||| ||||||||||||||| |||||||||||||
CCGGGCATCGCTCGCTGTCCGCCCAGGAGCCTGCCCAGGAGGAGCTGGTGGCAGAGGAG    249

GGAGCCCCCTGAACTGAATCCCCAGACAGAGAAAGCCAGGATGTGTGTACCTTTC    181
| ||| ||||||||||||||| |||||||||||||||||||||||||||||||
GGACCCGTCGGAACTGAATCCCCAGACAGAGAAAGCCAGGATCCTGCGCCCTTTC    309

ACAACTAGTCCGGCCTCGAAGAGTGCTTCCTAAAGGCCGGAAGGCGCGGCCTCGC    241
| |||||||| ||||||||| ||| || ||||||||||||||||| ||||||
CCGACTAGTTCGGCCTCGCAGAGTGCACCTAAAGGCCGGAAAAACACGGGCTCGA    369

TATTGCAGCCCATTATGAGGTTCATCCTCGGCCACCAGGACAGGATGGAGCACAAGCA    301
| ||||||||| |||||||| |||||||||||||||||||||||||||||||
GATCGCAGCCCATTATGAGGTTCATCCACGACCTGGACAGGAGGACGGAGCGCAGGCA    429

GGATGGGACAGTGAGTGGCTGGGAAGAGACCAAAATCAACAGCTCCAGCCCTCTG    361
|| ||||||||||||||| |||||||||||||||||||||| ||||||||||||
GGACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAGCCCTCTG    489

CGACCGGCCAGATTGGGAATTTACAGTCATCAGGCTGGCTCTCACTACTGTAC    421
| |||||| ||||||||||||||| ||||||||||||||||| |||||||||
CAACCGCCAGATCGGGAGTTTATAGTCACCCGGCGTGGCTCTACTACCTGTAC    549

GGTGCACTTTGATGAGGAAAGGCGTGTCACCTGAAGCTGGACTTGCTGGTGAAC    481
|||||||||||||||||||||| |||| ||||||||||||||||||||||||
GGTGCACTTTGATGAGGAAGGCGTGTCTACCTGAAGCTGGACTTGCTGGTGGAT    609

GCTGGCCCTGCGCTGCCCTGGAAGAATTCTCAGCCACAGCAGCAAGCTCTCCTGGG    541
|||||| ||||||||||||||| |||||||||||||||||||||||||||||
GCTGGCCCCTGCGCTGCCCTGGAATTCTCAGCCACTGCGGCCAGTTCCCTCGGG    669

GGTTCCGTTGTGCAGGTGTCTGGGCTGTGTTGCCGCTGGGCGGCCAGGGTCTCCCTT    601
|| ||||||| |||||||| |||||||||||||||||||||||||||||||||
GCTCCGCC TCTGCCAGGTGTCTGGGCTGTGTTGGCCCTGGGCCAGGGTCTCCCTG    729

CCGCACCCCTCCCCCTGGGCTCATCTTAAGGCTGCCCCCTTCCTAACCTACTTTGGA    661
|||||||||||||||||||||| ||||||||||||||||| |||||||||||||
CCGCACCCCTCCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCCACCTACTTCCGA    789

TCAAGTTCACTGAGGGGCCTTGCTTCTCCCAGATTCCTTAAACTTTCCCTGGCTCC    721
|||||||||||||| ||||||| |||||||||||| |||||||||||||||||
CCAGGTTCACTGAGGGGCCTGGTTCTCCCGCAGTCGTCCCAGGCTGCGGGCTCC    849

CATCACCACCTCCCTACCCCCACTCCTCGCACCCCCTC-GCTGCTCCTT    780
||||||||| |||| ||||||||||||||||||||||||
CCTGCACAGCTCTCTGGGCACCCGGTCCCCCTCTCGCCCACCCGTCTCTTT    904

AGTCCTGCTCCTCCC--TCAAAGGCAGCCAGGAGCTTGTTCACATGTTTCCATTCC-    837
|||| |||||||||||||||| |||||||||||||||||| ||||||
AGACCTGCCCCTCCCCTAGAGCGTTCGCCTGGGCCTTCACGTGTTTCCATCCC    964

--ACAGACGTATCCTTGCTCTTCTTAACATCCCATTCCACCACAACTATCCACCTC    891
||||||||||||||||||||||||||||||||||||||||||||||||||
AAATACAGTATTCCCACTCTTTATCTTACACACTCCCCCACCCCGCCCACTCTCCACCTC    1024

GCTCCCAAAGCCCCTAC-----------TTATCCCTGACTCCCCCACCCACT    936
|||||||| |||| |||| |||||||||
GCTCCCCAATCCTGACCCTTTGAGGCCCCAGTGATCTGACTCCCCCCTGGCCA    1084

CGACCACGTGTTTATTGACTTTGTGCAC-----------    968
||||| |||||||||||||||
CAGCCACCTGAGATGGCTTCCACTCTGTGGGCCAGGATGGTCATTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAG(
```

```
Db   1085  CAGACCCCCAGGTCATTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAG
           |||||| |||||||||||||||||||||||||||||||||||||| |
Qy    969  -----CAGGCACTGAGATGGGCTGGACCTGGTGCCAGGAGACCAGAGAACCTGGGA(
                 |||||||| |||||||||||| || |||||| ||||||| ||
Db   1145  CACTTCAGGCACTAAGAGGGGCTGGACCTGGCGCGGCCCGGCGGGGCCAAAGAGACTGGGC(
Qy   1024  GCCAGAAGTTCCCAACTGTGAGGGGGGCAAGAGCTGGGACAAGTCTCCCTGGA--
           |||||||||||||||||||||||||||||||| ||||||| ||||||| |
Db   1205  GCCAGGAGTTCCCAAATGTGAGGGGCGAGA-AACAAGACAAGCTCCTCCCTTGAGA
Qy   1080  CCTGTGGATTTTGAAA--AGATACTATTTTTATTATTATTGTGACAAAATGT---T
           ||||||||||||||| |||| |||||||||||||||||||||| |||
Db   1264  CCTGTGGATTTTTAAAACAGATATTATTTTATTATTATTGTGACAAAATGTTGAT
Qy   1135  GGATATTAAAGAGAATAAATCATGA    1159
           ||||||| |||||| ||||||||
Db   1324  GGATATTAAATAGAATAAGTCATAA    1348
```

RESULT 5
ABK40255
ID   ABK40255 standard; cDNA; 1353 BP.
XX
AC   ABK40255;
XX
DT   15-JUL-2002 (first entry)
XX
DE   cDNA encoding human PRO207 polypeptide.
XX
KW   Human; PRO; benign tumour; malignant tumour; lymphoid malignancy;
KW   leukaemia; neuronal disorder; stromal disorder; blastocoelic diso.
KW   inflammatory disorder; immune disorder; angiogenic disorder;
KW   gene therapy; cytostatic; neuroprotective; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153486-A1.
XX
PD   26-JUL-2001.
XX
PF   11-FEB-2000; 2000WO-US003565.
XX
PR   08-MAR-1999;   99WO-US005028.
PR   11-MAR-1999;   99US-0123972P.
PR   11-MAY-1999;   99US-0133459P.
PR   02-JUN-1999;   99WO-US012252.
PR   22-JUN-1999;   99US-0140650P.
PR   22-JUN-1999;   99US-0140653P.
PR   20-JUL-1999;   99US-0144758P.
PR   26-JUL-1999;   99US-0145698P.
PR   28-JUL-1999;   99US-0146222P.
PR   17-AUG-1999;   99US-0149395P.
PR   31-AUG-1999;   99US-0151689P.
PR   01-SEP-1999;   99WO-US020111.
PR   15-SEP-1999;   99WO-US021090.
PR   30-NOV-1999;   99WO-US028313.
PR   01-DEC-1999;   99WO-US028301.
PR   01-DEC-1999;   99WO-US028634.
PR   05-JAN-2000;   2000WO-US000219.
XX
PA   (GETH) GENENTECH INC.
XX
PI   Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;
PI   Marsters SA, Pan J, Pitti RM, Roy MA, Smith V, Stone DM;
PI   Watanabe CK, Wood WI;
XX
DR   WPI; 2002-205567/26.
DR   P-PSDB; AAU86129.
XX
PT   Thirty five nucleic acids encoding PRO polypeptides, useful for t
PT   benign or malignant tumors, leukemias and lymphoid malignancies,
PT   inflammatory, angiogenic and immunologic disorders.

3; 302pp; English.

nvention relates to the isolation of novel human PRO
and the polynucleotide sequences encoding them. The PRO
agonists, antagonists or anti-PRO antibodies are useful for
gn or malignant tumours (e.g. renal, kidney, bladder,
leukaemias and lymphoid malignancies, other disorders such
glial, astrocytal, hypothalamic, glandular, macrophagal,
blastocoelic disorders, inflammatory, immune and angiogenic
he polynucleotide sequences are also useful in gene therapy.
0288 encode for the human PRO polypeptides of the invention

3 BP; 257 A; 443 C; 389 G; 264 T; 0 U; 0 Other;

larity   53.8%;   Score 628.6;   DB 6;   Length 1353;
           76.2%;   Pred. No. 1.2e-159;
onservative   0;   Mismatches 219;   Indels 73;   Gaps 10;

[Sequence alignment blocks — left column]

Qy GAACAACTAGTCCGGCCTCGAAGAAGTGCTTCCTAAAGGSCCGGAAGGCCGGCCTCGC 241
Db AACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAACACGGGCTCGA 369

Qy GCTATTGCAGCCCATTATGAGTTCATCCTCGGCCAGGACAGGATGGAGCACAAGCA 301
Db CGATCGCAGCCCATTATGAAGTTCATCCAGACCTGGACAGGACAGGAGCCGCAGGCA 429

Qy GTGGATGGGAACAGTGAGTGGCTGGGAAGAGACCCAAAATCAACAGCTCCAGCCCTCTG 361
Db GTGGACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTTCAGCCCTCTG 489

Qy TACGACCGGCCAGATTGGGGAATTTACAGTCATCATCAGGGCTGGGCTCTACTACCTGTAC 421
Db TACAACCGCCAGATCGGGAGTTTATAGTCACCCGGGCTGGGCTTACTACTACCTGTAC 549

Qy CAGGTGCACTTTGATGAGGGAAAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGAAC 481
Db CAGGTGCACTTTGATGAGGGGAAAGGCCCTGTCTACCTGAAGCTGGACTTGCTGGTGGAT 609

Qy GTGCTGCCCTGCGCTGCGTGCTGGAAGAATTCTCAGCCACCAGACAGCTCTCCTGGG 541
Db GTGCTGGCCCTGCGCTGCGTGCTGGAGGAATTCTCAGCCCACTGCGCGCGAGTTCCCTCGGG 669

Qy CAGCTCCGTTGTGCCCAGGTGTCTGGGCGTTGCCGCTGCGGGCCAGGGTCTTCCCTT 601
Db CAGCTCCGCCCTGCGCCAGGTGTCTGGGCTGTTGGCCCTGCGGCGTCCCCCTG 729

Qy ATCCGACACCCTCCCCCTGGGCTCATCTTAAGGCTGCCCCTTCCTAACCTACTTTGGA 661
Db ATCCGACCACCCTCCCCCTGGGCCATCCAAGGCTGCCCCCTTCCTCACCTACTTCGGA 789

Qy TTTCAAGTTCACTGAGGGGCCTTGCTCCTCCCAGATTCCTTAAACTTTCCCTGCTCC 721
Db TTCCAGGTTCACTGAGGGGCCCCTGGTGTCCCCAGTCGTCCCAGGCTGCCGGCTCC 849

Qy AGCATCACCACACTCCCTACCCCCCACCCCCACTCCTCCACCCCCTC-GCTGCTCCTT 780
Db -CCTCGACAGCTCTCTGGGCACCCGGTCCCCTCTCCGCCCCACCCTCAGCCGCTCTTT 904

Qy CCCAGTTCCTGTCTCTCC--TCAAAGGCAGCCAGAGCTTGTTCACATGTTTCCATTCC- 837

[Sequence alignment blocks — right column]

Db 905 GCTCCAGACCTGCCCCTCCCTCGAGAGGCTGCCTGCTGGCCGCTGTTCACGTGTTTTCC
Qy 838 ------ACAGACGTATCCTGCTCTTCTTAACATCCACCACAACTATCC
Db 965 ACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCCACCGCCCACTCTCC

Qy 892 ACTAGCTCCCAAAGCCCTAC----------TTATCCTGACTCCCCAC
Db 1025 ACTAGCTCCCCAATCCCTGACCCTTGAGGCCCCCAGTGATCTGACTCCCCCT

Qy 937 CACCCGACCACCGTGTTTATTGACTTGTCAC---------
Db 1085 CAGACCCCCAGGTCATTGTGTTCACTGTACTCTTGTGGGCAAGGATGGGTCCAGAA

Qy 969 ------CAGGCACTGAGATGGGCTGGACCTGGTGGCAGGAAGCCAGAGAACCTGGG
Db 1145 CACTTCAGGCACTCAAGAGAGGGCTGGACCTGGCGGCGGCGGCCAGGAAGCCAAAGAGACTGGG

Qy 1024 GCCAGAAGTTCCCAACTGTGAGGGGGAAGAGCTGGGACAAGCTCCTCCCTGGA
Db 1205 GCCAGGAGTTCCCAAATGTGAGGGGCGAGA-AACAAGACAAGCTCCTCCCTTGAG

Qy 1080 CCTGTGGATTTTGAAA--AGATACTATTTTTATTATTATTGTGACAAAATGT--
Db 1264 CCTGTGGATTTTTAAAACAGATATTATTTTTTATTATTATTGTGACAAAATGTTGA

Qy 1135 GGATATTAAAGAGAATAAATCATGA 1159
Db 1324 GGATATAAATAGAATAAGTCATAA 1348

RESULT 6
AAX56000
ID   AAX56000 standard; DNA; 1421 BP.
XX
AC   AAX56000;
XX
DT   15-JUL-1999 (first entry)
XX
DE   Human tumour necrosis factor Apo-3 ligand polynucleotide sequenc
XX
KW   Human; tumour necrosis factor; Apo-3 ligand; lymphotoxin; apopto
KW   NF-kappaB-dependent transcription; JNK/SAPK-dependent response;
KW   ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   92..841
FT   /tag= a
FT   /product= "Apo-3 ligand"
XX
PN   WO9919490-A1.
XX
PD   22-APR-1999.
XX
PF   09-OCT-1998; 98WO-US021407.
XX
PR   10-OCT-1997; 97US-0062037P.
PR   17-DEC-1997; 97US-0069862P.
XX
PA   (GETH) GENENTECH INC.
XX
PI   Ashkenazi AJ, Marsters SA, Pitti R;
XX
DR   WPI; 1999-287982/24.
DR   P-PSDB; AAY09369.
XX
PT   New human Apo3- ligand (a tumor necrosis factor) homologue.
XX
PS   Claim 18; Fig 1; 74pp; English.
XX
CC   The present sequence encodes a human tumour necrosis factor (TNF

mologue designated Apo-3 ligand. Apo-3 ligand has
ivity. Apo-3 ligand can be used to induce apoptosis in
er cells, to induce NF-kappaB-dependent transcription and
SAPK-dependent responses in mammalian cells

BP; 281 A; 464 C; 404 G; 272 T; 0 U; 0 Other;

53.8%;   Score 628.6;   DB 2;   Length 1421;
rity   76.2%;   Pred. No. 1.2e-159;
nservative   0; Mismatches 219; Indels   73; Gaps   10;

```
GGAGGCTGGGGCTGGGGCTGGCCTGGCCTCCTGCTGGTGCGTGGTCAGCCTG    61
      ||||||| |||||||||||||||||||||||||||||||||||||
GGGGCTGGGGCTGGCGCTGGCCTGCCTCCTGCGCCTCCGCTCGTGGTCAGTTG    223

GCTGGGCAACGCTGTCTGCCCAGGAGCCTTCTCAGGAGGAGAGCTGACAGCAGAGGAC    121
      |||||||||||||||||||||||||||||||||||||||||||||||||||
GCCGGGCATGCTGTCGCCCCAGGAGCCTGCCCAGGAGGAGAGCTGGTGCGGCCAGGGAGGAG    283

GGGAGCCCCTGAACTGAATCCCCAGACACAGAGGAAAGCCAGGATGTGGTACCTTTC    181
     ||||||||||||||||||||||||||||||||||||||||||||| |||||
AGGACCCGTCGGAACTGAATCCCCAGACACAGAGGAAAGCCAGGATCCTGCGCCTTTC    343

ACAACTAGTCCGGCCTGGAAGAAGTGCTCCTAAAGGCCGAAGGCGCGGCCTCGC    241
     ||||||||||||||||||||||||| ||||||||||||||||||||||||
ACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAACACGGGCCTCGA    403

TATTGCAGCCCATTATGAGGTTCATCCTCGGCCAGGACAGGATGGAGCACAAGCA    301
     |||||||||||||||||||||||||||||||||||||||||||||||||
GATCGCAGCCCATTATGAAGTTCATCACGACCTTGGACAGGAGCGACGGAGGCCAGGCA    463

GGATGGGACAGTGAGTGGCTGGGAAGAGACCAAAATCAACAGTTCCAGCCCTCTG    361
     |||||||||||||||||||||||||||||||||||||||||||||||||
GGACGGGACAGTGAGTGGCTGGGAGGAAGCGCAGAATCAACAGTTCCAGCCCTCTG    523

ACGACCGCCAGATTGGGGAATTTACAGTCATCAGGGCTGGGCTTCTACTACCTGTAC    421
     |||||||||||||||||||||||||||| ||||||||||||||||||||||
CAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTTCTACTACCTGTAC    583

AGGTGCACTTTGATGAGGGAAAGGCTGTCTACCTGGAAGCTGGACTTGCTGGTGAAC    481
     |||||||||||||||||||||||||||| ||||||||||||||||||||||
AGGTGCACTTTGATGAGGGAAGGGCTGTCTACCTGGAAGCTGGACTTGCTGGTGGAT    643

GCTGCCCTGCGCTGCCTGGAAGAATTCTCAGCCACAGCAGCAAGCTCTTCCTGGG    541
     ||||||||||||||||||||| |||||||||||||||||||||||||||
GCTGGCCCTGCGCTGCCTGGAGGAATTCTCAGCCACCTGCGGGGGAGTTCCCTCGGG    703

AGCTCCGTTTGTGCCAGGTGTCTGGGCTGTTGCCGCGTGCGCGCGCAGGGTCTTCCCTT    601
     |||||||||||||||||||||||||||||||| ||||||||||||||||||||
AGCTCCGCCTCTGCCAGGTGTCTCGGGCTGTTGCGGCCTGTTGGCCCTGCGCGGCCAGGGTCTTCCCTG    763

CCGCACCCTCCCCCTGGGCTCATCTTAAGGCTGCCCCTTCCTAACCTACTTTGGA    661
     ||||||||||||||||||||||||||||||||||||||||||||||||||
CCGCACCCTCCCCCTGGGCCCCATCTCAAGGCTGCCCCCTTCCTCACCTACTTCGGA    823

TCAAGTTCACTGAGGGGCCTTGCTCTCCCAGATTCCTTAAACTTTCCCCTGGCCTCC    721
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCAGGTTCACTGAGGGGCCCCGGTCTCCCCGCAGCGTCCCCAGCGTGCCGGCCTCC    883

GCATCACCACACCTCCCTACCCCACCCCCACTCCTCCGCACCCCCTC-GCTGCTCCTT    780
     ||||| |||     |||||||||||| ||||| |||||||||||||
CCTGCGACAGTCTCTGGCCACCCGGTCCCCCCTCTGCCCACCCCGGCCGCTCTTT    938

CAGTCCTGCTGTCTCC--TCAAAGGCAGCCAGGCCGAGCTTGTTCACATGTTCCATTCC-    837
     |||||    |||  ||| |||||||||||| ||||||||||||||||||||
CAGACCTGCCCCCTCCCCTCTAGAGCGTTGCCTGGGCCTGTTTCACGTGTTTTCCATCCC    998

--ACAGACGTATCCTTGCTCTTCTTAACATCCCATCCACCACAACTATCCACCTC    891
     ||||||| ||| ||| |||| ||||||||||||||||||||||||||||
AAATACAGTATTCCCACTCTTATCTTACACTCCCCACCTCTCCACCTC    1058

GCTCCCAAAGCCCCTAC------------TTATCCCTGACTCCCCCACCCACT    936
|||||| |||  ||||                    ||||||||||||||
```

```
Db    1059  ACTAGCTCCCAATCCTGACCCCTTGAGGCCCCCAGTGATCTCGACTCCCCCTG

Qy     937  CACCCGACCACGTGTTTATTGACTTTGTGCAC---------------
             ||||||| ||||||||||||||||||||||
Db    1119  CAGACCCCCAGGTCATTGTGTTCACTGTACTCGTGGGCAGGAGGATGGGTCCAGAAG.

Qy     969  ----CAGGCACTGAGATGGGCTGGAGACCTGGTGGCAGGAAGCCAGAGAACCTGGGA
                   |||||||||||||||||||||||||||||||||||||||||||
Db    1179  CACTTCAGGCACTAAGAGGGCTGAGACCTGGCGCGGCAGGAGCCCAAAGAGACTGGGC

Qy    1024  GCCAGAAGTTCCCAACTGTGAGGGGGAAGAGTCGGGGACAAGCTCCTCCCTGGA-
             |||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    1239  GCCAGGAGTTCCCAAATGTGAGGGGCGAGA-AACAGACAAGCTCCTCCCTTGAGA

Qy    1080  CCTGTGGATTTGAAA--AGATACATTTTTATTATTGTGACAAAATGT---T.
             ||||||||||||||  ||||||||||| ||||||||||||||||||   |
Db    1298  CCTGTGGATTTTTAAAACAGATATTATTTTTTATTATTATTGTGACAAAATGTTGAT.

Qy    1135  GGATATTAAAGAGAATAAATCATGA  1159
             ||||||||| |||||||||||||||
Db    1358  GGATATTAAATAGAATAAGTCATAA  1382
```

RESULT 7
ACC57587
ID   ACC57587 standard; DNA; 1306 BP.
XX
AC   ACC57587;
XX
DT   28-JUL-2003   (first entry)
XX
DE   Polynucleotide encoding tumour necrosis factor superfamily member
XX
KW   Human; RANKL; tumour necrosis factor; osteopathic; bone; gene; ds
XX
OS   Homo sapiens.
XX
PN   WO200303663-A2.
XX
PD   24-APR-2003.
XX
PF   15-OCT-2002; 2002WO-US033022.
XX
PR   15-OCT-2001; 2001US-0329393P.
XX
PA   (BARN-) BARNES-JEWISH HOSPITAL.
XX
PI   Lam J, Ross PF, Teitelbaum SL;
XX
DR   WPI; 2003-430346/40.
XX
PT   New RANKL mimic comprising a core, and at least one external loop
PT   for enhancing processes of bone formation or inhibiting bone reso
PT   thus providing treatments for disease or condition characterized .
PT   of bone mass.
XX
PS   Disclosure; Page 66-67; 78pp; English.
XX
CC   The present sequence is that of a polynucleotide encoding a non-R
CC   member of the tumour necrosis factor (TNF) superfamily. The inven
CC   provides non-naturally-occurring proteins that contain one or mor
CC   external surface loops of RANKL (see ABR42066-70) in combination
CC   heterologous protein core obtained from a non-RANKL member of the
CC   superfamily. Also provided are polynucleotides encoding such prot
CC   The proteins bind to RANK, acting as mimics of RANKL. They can be
CC   enhance bone formation by either inhibiting bone resorption or in
CC   osteogenesis, thus providing treatment for diseases or conditions
CC   characterised by loss of bone mass
XX
SQ   Sequence 1306 BP; 247 A; 434 C; 368 G; 257 T; 0 U; 0 Other;

Query Match   53.4%;   Score 624;   DB 7;   Length 1306;
Best Local Similarity   76.0%;   Pred. No. 2.1e-158;

Conservative　0;　Mismatches　220;　Indels　73;　Gaps　10;

```
Qy    TGAGCCTGGGCCTGGCGTGGCCTTGGCCTTCGTCCTGCTCCTGCTGTCTGTGTCAGCCTG    61
         |||||||||||||||||||||||| |||||||||||||||| |||||||
Db    TGGGCCTGGGCCTGGCGTGGCCTGGCCTTCGTCCTGCTCCTGCTGTCTGGTCAGTTTG    149

Qy    AGCTGGGCAACGTGTCTGCCCGCCGCAGAGCCTTCAGGAGGAGCTGACAGCAGAGGAC    121
         ||||||||||||| || |||||||||||||||||||||||||||| |||||
Db    GCCCGGGCATCGCTGTCCGCCCAGAGCCTGCCGCCAGGAGCAGAGGAGCTGGTGCAGAGGAG    209

Qy    GGGAGCCCCCTGAACTGACTCCCAGACAGAGGAAAGCCAGGATGTGGTACTTTC    181
         |||||||  ||||| |||||||||||||||||||||||| ||||||||
Db    AGGACCCGTCGGAACTGAATCCCAGACAGAAGAGAAAGCCAGGATCCTGCGCCTTTC    269

Qy    AACAACTAGTCCGCCTCGAAGAAGTGCTCCTAAAGGCCGGAGGCGCGGCCTCGC    241
         ||||||||| || ||||||||||||||||||||||||||||
Db    ACCGACTAGTTCGCCTCCCAGAGAGTGCACCTAAAGGCCGGAAAACACGGGCTCGA    329

Qy    CTATTGCAGCCCATTATGAGGTTCATCCTCGGCCAGGACAGGATGAGCACAAGCA    301
         ||||||||  ||||||||||||||||||||||||||
Db    CGATTCGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACAGGCAGCAGCA    389

Qy    TGGATGGACAGTGAGTGCTGGGAAGAGACCAAAATCAACAGCTCCAGCCCCTG    361
         ||||||||||||||||||||||||  ||||||||||||||||||||
Db    TTGGACGGGACAGTGAGTGCTGGGAGGAGAATCAACAGCTCCAGCCCCTG    449

Qy    ACGACCGGCCAGATTGGGGAATTTACAGTCATCAGGGCTGGGCTCTACTACCTGTAC    421
         |||||||| |||||||| ||||||||||||||||||||||||||||||||
Db    ACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTACTACCTGTAC    509

Qy    AGGTGCACTTTGATGAGGGAAAAGCTGTCTACTGGACTGGTGTTGCTGGTGAAC    481
         |||||||||||||||||||||||||||||||||||||||||
Db    AGGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGGTGAT    569

Qy    TGCTGGCCCTGCGCTGCCTGGAAGAATTCTCAGCCACCAGCAGCAAGCTCCTGGG    541
         ||||||||||||||||||||||||||||||||||||||||||||
Db    TGCTGGCCCTGCGCTGCCTGGAGAATTCTCAGCCACTGCGCGGCCAGTTCCCTCGGG    629

Qy    AGCTCCGTTGTGTGCAGGTGTCTGGGCTGTTGCCGTCGGGCCAGGGTCTTCCCTT    601
         ||||||||||||||||||||||||||||||||||||||||||
Db    AGTTCCGCCTCTGCCTGCGGTGTCTGGGCTGTGGGCTGCCCTGGGGTCTCCTCCCTG    689

Qy    ATCCGCACCCTCCCCTGGGGCTCATCTTAAGGCTGCCCCTTCCTAACCTACTTTGGA    661
         ||||||||||||||||||||||||||||||||||||||||
Db    TCCGGCACCCCTCCCCCGGGCCCCATCCCAAGGCTGCCCCTTCCTCACCTACTTCGGA    749

Qy    TTCAAGTTCACTGAGGGGGCCTTGCTCTGCTCCCAGATTCCTTAAACTTTCCCTGGCTCC    721
         ||| ||||||||||||||||||||||| ||||||||||||||
Db    TCCAAGTTCACTGAGGAGGGCCCTGGGCCCTCCTCCCCACAGTCGTCCCAGGCTGCCGGTCC    809

Qy    AGCATCACCACACCTCCTACCCCACCCCCACTCCTCCACCCCTC-GCTTGCTCCTT    780
         ||||||||  ||||||||||||||||||||||||||
Db    TAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCACCTCTCCACCTCAGCGCTCTT    864

Qy    --CCTCGACAGCTCCTGGGCACCCGGTCCCCGGTCCCCTGCCCCACCCCACCCGGCTCTTT    837
         |||||||||||||||||||||||||||||||||
Db    CCAGTCCTGCTCTCTC--TCAAAGGCAGCCAGCAGGTTGTTCACATGTTCCATTCCATCCC    924

Qy    CCAGACCTGCCCCTCCCTCCTTAGAGGCTGCCTGGGCCTGTTCACGTGTTTTCCATCCC    891
         --ACAGACGTATCCTTGCTCTTCTTAACATCCATCCTTAACATCCAACAACTATCCACCTC    984

Qy    AGCTCCCAAGCCCCTAC-----------TTATCCCTGACTCCCCACCCACT    936
         ||||||||  ||||||||||
Db    AGCTCCCCAATCCCTGACCCTTTGAGGCCCCCAGTCATCTGACTCCCCCCCTGGCCA    1044

Qy                                                       968
Db    CCGACCACGTGTTTATTGACTTTGTGCAC----------    1104

Qy    ACCCCCAGGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAGAGACCC    1104
         --CAGGCACTGAGATGGGCCTGACCTGGTGCACAGAAGCCAGGAAGCCAGAGAACCTGGGACTAG    1023
Db    TTCAGGCACTAAGAGAGGGGCTGGACCTGGCGGCCAGGAAGCCAAAGAGACTGGGCCTAG    1164

Qy    1024   GCCAGAAGTTCCCAACTGTGAGGGGGAAGAGAGCTGGGACAACAAGCTCCTCCCTGGA--
              ||||||||||||||||||||||||  ||||||||||
Db    1165   GCCAGGAGTTCCCAAATGTGAGGGGCCGAGA-AACAAGACAAGCTCCTCCCTTGAGA

Qy    1080   CCTGTGGATTTGAAA--AGATACTATTTTATTATTATTGACAAAATGT---]
              ||||||||||||||  |||||||||||||||||||||||||||
Db    1224   CCTGTGGATTTTTAAAACAGATATTATTTTTATTATTATTGTGACAAAATGTTGAT

Qy    1135   GGATATTAAAGAGAATAAATCA    1156
              ||||||||| ||||||| |||
Db    1284   GGATATTAAATAGAATAAGTCA    1305
```

RESULT 8
ACC57901
ID　ACC57901 standard; cDNA; 1306 BP.
XX
AC　ACC57901;
XX
DT　11-AUG-2003　(first entry)
XX
DE　Human TWEAK coding sequence.
XX
KW　Human; TWEAK; tumour necrosis factor; ligand; cytostatic;
KW　immunomodulator; osteopathic; gene; ss.
XX
OS　Homo sapiens.
XX
FH　Key　　　　Location/Qualifiers
FT　CDS　　　　18..767
FT　　　　　　 /*tag= a
FT　　　　　　 /product= "Human TWEAK"
XX
PN　WO2003040307-A2.
XX
PD　15-MAY-2003.
XX
PF　25-JUL-2002; 2002WO-US023782.
XX
PR　27-JUL-2001; 2001US-0307838P.
XX
PA　(HUMA-) HUMAN GENOME SCI INC.
XX
PI　Hilbert DH,　Rosen CA;
XX
DR　WPI, 2003-430659/40.
DR　P-PSDB; ABR42315.
XX
PT　New heteromultimeric complex having a first polypeptide member o
PT　tumor necrosis factor (TNF) ligand family, and a second differen
PT　of TNF ligand family, useful for treating cancer, osteoporosis o
PT　autoimmune disease.
XX
PS　Disclosure; Page 367-368; 388pp; English.
XX
CC　The present sequence is that of a polynucleotide encoding human '
CC　The invention relates to compositions comprising heterotrimeric (
CC　of tumour necrosis factor (TNF) ligand family members, and their
CC　the detection, prevention and treatment of disease. In one embod
CC　the heterotrimeric complex comprises full-length or extracellula
CC　portions of TWEAK and full-length or extracellular portions of o
CC　ligand family members, preferably VEGI or VEGI-SV. The heterotri
CC　complexes of the invention are useful for treating an autoimmune
CC　cancer or osteoporosis, and particularly for inhibiting cancer c
CC　proliferation, increasing B cell proliferation, or inducing apop
CC　T cells
XX
SQ　Sequence 1306 BP; 247 A; 434 C; 368 G; 257 T; 0 U; 0 Other;

Query Match　　　　 53.4%;　　Score 624;　DB 7;　Length 1306;
Best Local Similarity　76.0%;　Pred. No. 2.1e-158;
Matches 929;　Conservative　0;　Mismatches 220;　Indels　73;

```
Qy   1024  GCCAGAAGTTCCCAACTGTGAGGGGGAAGAGCTGGGACAAGCTCCTCCTGGA--
                 ||| ||||||||||||||||||||||||||||||   |||||||||||||||
Db   1165  GCCAGGAGTTCCCAAATGTGAGGGGCGAGA-AACAAGACAAGCTCCTCCCTGAGA

Qy   1080  CCTGTGGATTTTGAAA--AGATACTATTTTTATTATTATTGTGACAAAATGT---T.
                 ||||||||||||||||   ||||||||| ||||  |||||||||||||||||    |
Db   1224  CCTGTGGATTTTTAAAACAGATATTATTTTTTATTATTATTGTGACAAAATGTTGAT.

Qy   1135  GGATATTAAAGAGAATAAATCA  1156
                 ||||||||| ||||||||||||
Db   1284  GGATATTAAATAGAATAAGTCA  1305
```

RESULT 9

```
ADC35205
ID   ADC35205   standard; cDNA; 1306 BP.
XX
AC   ADC35205;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Human cDNA encoding TNF ligand family member #12.
XX
KW   ss; gene; human; tumour necrosis factor; TNF ligand; endokine alp
KW   excessive bone resorption disorder; osteoporosis; Paget's disease
KW   arterial calcification.
XX
OS   Homo sapiens.
XX
PN   US2003100074-A1.
XX
PD   29-MAY-2003.
XX
PF   15-AUG-2002; 2002US-00218547.
XX
PR   16-AUG-2001; 2001US-0312542P.
PR   30-OCT-2001; 2001US-0330761P.
XX
PA   (YUGG/)  YU G.
PA   (NIJJ/)  NI J.
PA   (ROSE/)  ROSEN C A.
PA   (NARD/)  NARDELLI B.
XX
PI   Yu G,  Ni J,  Rosen CA,  Nardelli B;
XX
DR   WPI; 2003-696072/66.
DR   P-PSDB; ADC35206.
XX
PT   New Endokine alpha gene useful for preparing a composition for tr
PT   disease associated with excessive or insufficient bone resorption
PT   osteoporosis, Paget's disease or arterial calcification.
XX
PS   Disclosure; SEQ ID NO 23; 145pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule encodi
CC   tumour necrosis factor family ligand. A composition comprising th
CC   isolated antibody or its fragment is used for treating an individ
CC   need of decreased level of endokine alpha activity. The endokine
CC   polypeptide present in a heterotrimeric complex is used for treat
CC   individual having a disorder associated with excessive bone resor
CC   e.g. osteoporosis, Paget's disease or arterial calcification. Tre
CC   individual having a disorder associated with insufficient bone re
CC   comprises administering an endokine alpha antagonist, which is th
CC   antibody that binds specifically to endokine alpha polypeptide. T
CC   present sequence represents a cDNA encoding a tumour necrosis fac
CC   family ligand.
XX
SQ   Sequence 1306 BP; 247 A; 434 C; 368 G; 257 T; 0 U; 0 Other;

Query Match             53.4%;  Score 624;  DB 9;  Length 1306;
Best Local Similarity   76.0%;  Pred. No. 2.1e-158;
Matches  929;  Conservative    0;  Mismatches  220;  Indels   73;   G
```

```
Qy   ?GAGCCCTGGCCTGGCGCTGGCCTCTGCCTCTGCTGCTGGTCGTGGTCAGCCTG      61
           ||||||||||||| |||||||||||||| ||||||||| ||||||||||||||
Db   ?GGGCCTGGGCCTGGCCTCGCCCTGCCTCCGCCTCCGCTGCCGTGGTCAGTTTG     149

Qy   ?CTGGGCAACGCTGCTGCCGCCAGGAGCCTTCTCAGGAGGAGCTGACAGCAGAGAGAC    121
           ||||| | |||| ||||||||||| |||| ||| |||||||||||||||||||| | |
Db   ?CCCGGGCATCGCTGTTCCGCCCAGGAGCCTGCCCAGGAGGAGCTGTGGCAGAGGAG    209

Qy   ?GGAGCCCCCTGAACTGAATCCCAGACAGAGAAGCCAGGATGTGGTACCTTTC       181
           ||| ||| |||| |||||||||||| ||||||||||| |||||||||| ||||
Db   ?GGACCCGGTGGAACTGAATCCCAGACAGAGAAGCCAGGATCCTCGGCCTTTC      269

Qy   ?ACAACTAGTCCGGCCTCGAAGAAGTGTCTCTAAAGGCCCGGAAGGCGCGGCCTCGC    241
           || | || |||| |||||||||| ||||| |||||||||||| |||||| ||||||
Db   ?ACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAAACACGGGCCTCGA   329

Qy   ?CTATTGCAGCCCATATGAGGTTCATCCTCGGCCCAGGACAGGATGAGACAAGCA      301
           |  | |||||||| ||||| |||||| |||| |||||| ||||||||| ||| ||
Db   ?CGATGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACGGAGCCCAGGCA    389

Qy   ?TGGATGGGACAGTGAGTGGCTGGAAGAGACAACCAAAATCAACAGCTTCCAGCCCTCTG    361
           ||| || |||||| |||| |||||| | || ||| |||| ||||| ||||||||||||
Db   ?TGGACGGGACAGTGAGTGGCTGGGGAGGAAGCCAGAATCAACAGCTTCCAGCCCTCTG   449

Qy   ?ACGACGCCAGATTGGGGAATTTACAGTCATCAGGGCTGGGCTCTACTACCTGTAC      421
           || ||| ||||   |||||||||| |||||||| ||| |||||||||||||||||
Db   ?ACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTACTACCTGTAC    509

Qy   ?AGGTGCACTTTGATGAGGGAAAGGCTGTCTACCTGGACTTGCTGGTGAAC        481
           ||||||||||||||||| |||||||||| |||||||||||| ||||| | ||
Db   ?GGTGCACTTTGATGAGGGGAAGGCTGTGTCTACCTGGACTTGCTGGTGGAT      569

Qy   ?CGTGGCCCTGCGCTGCCTGGAAGAATTCTCAGCCACAGCAGCAAGCTCTCCTGGG    541
           || ||||| || || |||||| |||| ||| ||||||||| ||||||| ||||||
Db   ?GGCTGGCCCTGCGCTGCCTCTGAGGAATTCTCAGCACCACTGCGCGGCCAGTTCCCTCGGG   629

Qy   ?AGCTCCGTTTGTGCCAGGTGTCTGGGCTGTGCCGCTGGCGGCCAGGGTCTTCCCTT    601
           ||||||| ||||||||||| || || || ||||| ||||| ||||| ||||||||
Db   ?AGCTCCGCCTGCCTGCCAGGTGTCTGGGCTGTTGGCCTGTTGGCCTGGGTCCTCCCCTG   689

Qy   ?TCCGCACCCTCCCCTGGGCTCATCTTAAGGCTGCCCCCTTCCTAACCTACTTTGA    661
           |||||| | ||||| ||| || ||| ||||||||||| ||| || ||||||| |||
Db   ?CCGCACCCTCCCCTGGGCCATCTCAGGCTGCCCCCTTCCTTCTCCTACTTTCGGA     749

Qy   ?TTCAAGTTCACTGAGGGCCTTGCTCTCCCAGATTCCTAAAACTTTCCCTGGCTCC    721
           |||||||||||||||||||||||  |||||||  | ||||| |||| ||||||||
Db   ?TCCAGGTTCACTGAGGGGCCCTGGTCTCCCCCACAGTCGTCCCAGGTGCCGGCTCC   809

Qy   ?GCATCACCACACTCCCTACCCCACCCCACTCCTCCACCCCCTC-GCTGTCCTT     780
           ||| ||||||  | | ||| | ||||||||||| |||||||| | |||| ||||
Db   ?CCTGACAGCTCTCTGGCACCCGGTCCCCTGTCCCCCACCCCTCAGCCGCGTCTTT    864

Qy   ?CAGTCCTGTCCTCC--TCAAAGGCAGCCAGAGCTTGTTCACATGTTTCATTCC-    837
           ||||| | |||||    |||||||| ||||| ||| | ||||||||||||||||
Db   ?CAGACCTGCCCCTCCCTAGAGGCTGCCTGGGCTGTCGTCACGTGTGTTTCCATCCC    924

Qy   ?--ACAGACGTATCCTTGCTCTTCTTAACATCCCATGCCACCACAACTATCCACCTC     891
             || | |||| |||  | ||| ||||||||||||| ||||||||||||||||| |
Db   ?AAATACAGTATTCCCACTCTTATCTTACAACTCCCCCACCCGCCACTCTCCACCTC    984

Qy   ?GCTCCCAAAGCCCCTAC---------TTATCCCTGACTCCCCCACCCACT     936
           |||||| || ||||| ||         ||| ||| |||| ||||||| |||
Db   ?GCTCCCCAATCCCTGACCCCTTTGAGGCCCCCAGTGATCTTGCAGTCGATCTCGACTCCCCCTGGCCA   1044

Qy   ?CGACCAGGTGTTTATTGACTTGTGTGCAC--------     968
           | |||| || ||| || || | || ||||
Db   ?CCCCAGGGCATTGTGTTCACTGTACTCTGTGTGGGCAAGGATGGGTCCAGAGACCCC   1104

Qy   ?-CAGGCACTGAGATGGGGCTGACCTGGTGGCGAAGGAAGCCAGAGAACCTGGGACTAG   1023
            |||||||| | | | ||| |||||||||||| ||||||||||| || | |||| ||||
Db   ?TCAGGCACTAAGAGGGGCTGGACCTGGCGGAGCGGCGCAGGAAGCCAAAGAGACTGGGCCTAG   1164      G
```

Db   1165 GCCAGGAGTTCCCAAATGTGAGGGGCGAGA-AACAAGACAAGCTCCTCCCTTGAG.

Qy   1080 CCTGTGGATTTTGAAA--AGATACTATTTTATTATTATTGTGACAAAATGT---'

Db   1224 CCTGTGGATTTTTAAAACAGATATATTATTTTTATTATTATTGTGACAAAATGTTGA'

Qy   1135 GGATATTAAAGAGAATAAATCA 1156

Db   1284 GGATATTAAATAGAATAAGTCA 1305

RESULT 10
ABK34881
ID   ABK34881 standard; cDNA; 1364 BP.
AC
XX
XX
DT   08-MAY-2002 (first entry)
XX
DE   Human cDNA encoding secreted protein #19.
XX
KW   Human; secreted protein; gene; ss; nutritional supplement; haemo
KW   viral infection; bacterial infection; fungal infection; diabetes
KW   autoimmune disorder; rheumatoid arthritis; multiple sclerosis; t
KW   autoimmune thyroiditis; allergic reaction; neurodegenerative dis
KW   Alzheimer's disease; Parkinson's disease; liver fibrosis; cancer
KW   coagulation disorder; inflammatory disorder; Crohn's disease; in
KW   tissue regeneration; wound healing; burn; haematopoiesis;
KW   myeloid cell deficiency; lymphoid cell deficiency.
XX
OS   Homo sapiens.
XX
PN   WO200177288-A2.
XX
PD   18-OCT-2001.
XX
PF   29-MAR-2001; 2001WO-US010224.
XX
PR   06-APR-2000; 2000US-0195582P.
XX
PA   (GEMY) GENETICS INST INC.
XX
PI   Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnic
PI   Gulukota K, Graham JR;
XX
DR   WPI; 2002-179321/23.
XX
PT   Five hundred and ninety two polynucleotides derived from a varie
PT   human tissue sources which encode secreted proteins, useful for
PT   immune deficiencies and disorders such as autoimmune disorders.
XX
PS   Claim 1; Page 82; 372pp; English.
XX
CC   The invention relates to 592 polynucleotides which have been der
CC   a variety of human tissue sources and which encode novel secrete
CC   proteins. The polynucleotides can be used as probes for the
CC   identification and isolation of full length cDNA and genomic DNA
CC   polynucleotides and proteins can also be used as nutritional sup
CC   The proteins are useful in the treatment of various immune defic
CC   and disorders such as viral infections, bacterial infections, fu
CC   infections, autoimmune disorders (e.g. rheumatoid arthritis, mul
CC   sclerosis, autoimmune thyroiditis and diabetes) and allergic rea
CC   and conditions (e.g. asthma). They are also useful for treating
CC   neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson'
CC   disease; liver fibrosis, coagulation disorders (e.g. haemophili
CC   inflammatory disorders (e.g. Crohn's disease) and tumours. They
CC   useful for tissue regeneration, for wound healing and in the tre
CC   burns, incisions and ulcers. The proteins are also useful for re
CC   haematopoiesis and for treating myeloid or lymphoid cell deficie
CC   Sequences ABK34863-ABK35454 represent polynucleotides of the inv
XX
SQ   Sequence 1364 BP; 246 A; 461 C; 394 G; 263 T; 0 U; 0 Other;

```
          52.9%;  Score 618.2;  DB 6;  Length 1364;
rity      76.0%;  Pred. No. 7.8e-157;
nservative    0;  Mismatches 218;    Indels    73;    Gaps    10;

      GAGCCTGGGCGCTGGGCGCTGCCGCCTTGGCCTCGTCGTGGTCGTGGTCAGCCTG    61
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
      GGGCCTGGGCCTGGGCGCTGGGCGCCTCGGCCTTCGGCCGTGGTCAGTTTG       217

      CTGGGCAACGCTGTCTGCCCAGGAGCCTTCTCAGGAGGAGCTGACAGCAGAGGAC    121
      ||||||||| ||||||||| |||||||||||||||| ||||||||| |||||||||
      CCGGGCATCGCTGCGTCCCCAGGAGCCTGCCCAGGAGCAGCGGTGGTGGCAGAGGAG  277

      GGAGCCCCCTGAACTGAATCCCCAGACAGAGGAAAGCCAGGATGTGTTACCTTTC    181
      ||||||| |||||||||||||||||||||||||||||||||||||||| |||||||
      GGACCCGTCGGAACTGAATCCCCAGACAGACGAAGAGAAGCCAGGATCCTGCGCCTTTC  337

      ACAACTAGTCCGGCCCTGAAGAAGTGCTCCTAAAGGCCGGAAGGCGCGGCCTCGC    241
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      CCGACTAGTTCCGGCCTGCAGAAGTGCACCTAAAGGCCGGAAAACACGGGCTTCGA   397

      TATTGCAGCCCATTATGAGGTTCATCCTCGGCCAGACACGGATGGAGCACAAGCA    301
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      GATCGCAGCCCATTATGAGGTTCATCCACGACCTGACACGGACGAGCGCAGGCA   457

      GGATGGGACAGTGAGTGGCTGGGAAGAGACCAAAATCAACAGCTCCAGCCCTCTG    361
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      GGACGGGACAGTGAGTGGCTGGGAGAAGCCAGAATCAACAGCTCCAGCCCCTCTG   517

      CGACCGCCAGATTGGGGAATTTACAGTCATCAGGGCTGGGGCTCTACTACCTGTAC    421
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      CAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGGCTGGGCTCTACTACCTGTAC   577

      GGTGCACTTTGATGAGGGAAAGGCTGTCTACCTGAAGCTGACTTGCTGGTGAAC    481
      |||||||||||||||||||||||||||||||||| |||||||||||||||||||| |
      GGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAAAGCTGGACTTGCTGGTGGAT   637

      GCTGGGCCCTGCGCTGCTGGAGAATTCTCAGCCACAGCAGCAAGCTCTCCTGGG    541
      |||||||||||||||||||||||||||||||||||||||||||||||||||||| |
      GCTGGGCCCTGCGCTGCCTGGAGAATTCTCAGCCACTCGGCCCAGTTCCCTCCGG   697

      GCTCCGTTTGTGCCAGGTGTCTGGCGTTGCGCTGCGCGCCAGCGGTCTTCCCTT    601
      |||||||||||||||||||||||||||||| |||||||||||||||||||||| ||
      GCTCCGCCTCTGCCAGGTGTCTGGCGTGTTGGCCCTGCGCGCCAGGGGTCCTCCCTG   757

      CCGCACCCTCCCCGGGCTCATCTTAAGGCTGCCCCCTTCCTAACCTACTTTCGA    661
      |||||||||||||||||||||||||||||||||||||||||||||||||||||| |
      CCGCACCCCCCCCGGGCCCATCTCAAGGCTGCCCCCTTCCTCACCTACTTCCGGA   817

      TCAAGTTCACTGAGGGGCCTTGCTCTCCCAGATTCCTTAAACTTTCCCTGGCTCC    721
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
      CCAGGTTCACTGAGGGGCCCTGGTCTCCCCGCAGTCGTCCCAGCGTGCGGCTCC   877

      CATCACCACACCTCCCTACCCACCCCCACCCTCCTCCACCCCCTC-GCTGCTCCTT    780
      ||||||||||||||||||||||||||| |||||||||||||| |||||||| |||
      CCTCGACAGCTCTCTGGGCACCCGGTCCCCTGCCCCACCCCTCAGCCCGCTCTTT   932

      AGTCCTGCTCTCCC--TCAAAGGCAGCCAGAGAGCTTGTTCACAATGTTCCATTCC-   837
      |||||||||||||      ||||| |||||||||| ||||||||||||||||||||
      AGACCTGCCCCCTCCCTCTAGAGGCTGCCTGGGCCTGCTGGGGCCTGTTTTTTCCATCCG   992

      -ACAGACGTATCCTTGCTCTTCTTAAACATCCCATCCCACCCACAACTATCCACCTC   891
      ||||||||| |||| |||||||| |||||||||| ||||||||||||||||
      AATACAGTATTCCCACTCCTTATCTTACAACTCCCCACCGGCCCACTCTTCCACCTC   1052

      CTCCCAAAGCCCTAC----------------TTATCCCTGACTCCCCCACCCACT    936
      |||||||||||              |||||||||||||||||||||
      CTCCCCAATCCCTGACCCCTTGAGGCCCCAGTGATCTGACTCCCCCCTGGCCA   1112

      GACCCACGTGTTTATTGACTTTGTGCAC------------------    968
      ||||||      ||| ||| |||||||
      CCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGGTCCAGAAGACCCC  1172
```

```
          Qy  969  ------CAGGCACTGAGATGGGCTGGACCTGGTGCGCAGGAAGCCAGAGAACCTGGGAC
                         || ||||| |||| |||||||||||||| ||||||||||||| |||  |||||
          Db 1173  CACTTCAGGCACTAAGAGGGCTGGACCTGGCCTGGACCTGGCGCAGGAAGCAGGAAGCCAAAGAGACTGGGCC

          Qy 1024  GCCAGAAGTTCCCAACTGTGAGGGGGGAAGAGCTGGGGACAAGCTCCTCCCTGGA---
                    ||||||||||||||||||||||||||||||| ||||||| |||||||||||||||||
          Db 1233  GCCAGGAGTTCCCAAATGTGAGGGGCCGAGA-AACAAGACAAGCTCCTCCCTTGAGAA

          Qy 1080  CCTGTGGATTTTGAAA--AGATACTATTTTATTATTATTGTGACAAAATGT--TA
                    |||||||| |||||  |||  ||||||||||||||||||||||||||||||||||
          Db 1292  CCTGTGGATTTTTAAAACAGATATTATTTTATTATTATTGTGACAAAATGTGATA

          Qy 1135  GGATATTAAAGAG 1147
                    |||||||| ||||
          Db 1352  GGATATTAAATAG 1364
```

RESULT 11

AAV18600
```
ID   AAV18600 standard; cDNA; 1373 BP.
XX
AC   AAV18600;
XX
DT   21-JUL-1998  (first entry)
XX
DE   Homo sapiens tumour necrosis factor related ligand (TRELL) gene.
XX
KW   TRELL; tumour necrosis factor related ligand; tnf; treatment; canc
KW   autoimmune disease; immune system; stimulation; suppression;
KW   graft rejection; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..852
FT                   /*tag=   a
FT                   /note=  "tumour necrosis factor related ligand"
XX
PN   WO9805783-A1.
XX
PD   12-FEB-1998.
XX
PF   07-AUG-1997;  97WO-US013945.
XX
PR   07-AUG-1996;  96US-0023541P.
PR   18-OCT-1996;  96US-0028515P.
PR   18-MAR-1997;  97US-0040820P.
XX
PA   (BIOJ )  BIOGEN INC.
PA   (UYGE-)  UNIV GENEVA FACULTY MEDICINE.
XX
PI   Chicheportiche Y,    Browning JL;
XX
DR   WPI; 1998-145619/13.
DR   P-PSDB; AAW47525.
XX
PT   Tumour necrosis factor related ligand - useful for, e.g. treating
PT   auto-immune disease and immune responses to tissue grafts.
XX
PS   Claim 2; Page 48-50; 69pp; English.
XX
CC   The sequence is that encoding human tumour necrosis factor relate
CC   (TRELL). TRELL or active fragments can be included with a carrier
CC   pharmaceutical compositions to treat cancer, autoimmune diseases
CC   immune responses to tissue grafts, or to stimulate or suppress th
CC   system. It is useful to screen for TRELL receptors, by labelling
CC   detectable label and screening compositions for binding. Agents
CC   interfering with TRELL-receptor binding can also be screened for,
CC   then be administered, optionally with interferon- gamma, to induc
CC   death or treat, suppress or alter immune responses (especially in
CC   human adenocarcinoma cells) involving a signal pathway between TRI
CC   its receptor. The DNA sequence can be used in gene therapy for TRI
CC   related disorders in mammals (especially humans), e.g. tumours,
```

a inflammatory diseases or inherited genetic disorders, by
nto cells, and expressing, therapeutically effective amounts
e.g. a virus comprising a gene encoding TRELL. It may also
the preparation of prepare probes for screening
atic DNAs for TRELL-encoding sequences and for antisense

BP; 247 A; 462 C; 394 G; 270 T; 0 U; 0 Other;

52.6%; Score 614.6; DB 2; Length 1373;
arity   75.8%; Pred. No. 7.4e-156;
onservative   0; Mismatches 219; Indels 70; Gaps   9;

```
Qy  TGAGCCTGGGCCTGGCGCTGGCCTGCCTTGGCCTCCTGCTCGTGGTCAGCCTG     61
        ||||||||||||||||||||||||||||||||| ||||||||||| |||
Db  TGGGCCTGGGCCTGGCGCTGGCCTGCCTCCGGCCTCCTGCTGGTCAGTTTG    237

Qy  GCTGGGCAAGCGTGTCTGCCCAGGAGCCTTCTGCAGGAGGAGCTGACAGCAGAGGAC    121
        |||||||| |||||                   |||||||||||||||||||
Db  GCCGGGCATCGCTGTCCGCCCAGGAGCCTGCCCAGGAGGAGCTGGTGGCAGAGGAG    297

Qy  GGGAGCCCCCTGAACTGAACTCCCAGACAGAGGAAAGCCAGGATGTGGTACCTTTC    181
        ||| ||||||||                  ||||||||||||||||||||||
Db  AGGACCCCGTCGGAACTGGAATCCCAGACAGAGGAAAGCCAGGATCCTGCGCCTTTC    357

Qy  AACAACTAGTCCGGCCTCGAAGAAGTGCTCCTAAAGGCCGGAAGGCCGCGGCCTCGC    241
        ||||||||||||||||||||              |||||||||||||||||||
Db  ACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAAACAGGGGCTCGA    417

Qy  CTATTGCAGCCCATTATGAGGTTCATCCTCGGCCAGGACAGGATGGGAGCACAAGCA    301
        |||||||||||||||||||||||||||||  |||||||||||||||||||||||
Db  CGATCGCAGCCCATTATGAGGTTCATCCAGCGACCTGGACAGGACGGAGCGCAGGCA    477

Qy  TTGGATGGGACAGTGAGTGCTGGGAAGAGACCAAAAATCAACAAGCTCCAGCCCTCTG    361
        ||||||||||||||||||||||||||||||||        ||||||||||||||
Db  CTGGACGGGACAGTGAGTGCTGGGAAGAGACCAGAATCAACAAGCTCCAGCCCCTCTG    537

Qy  ACGACCGGCCAGATTGGGGAATTCACAGTCATCAGGGGCTGGGGCTTCTACACCTGTAC    421
        |||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db  ACAACCGGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGCTTCTACTACCTTGTAC    597

Qy  AGGTGCACTTGATGAGGGAAAGGCTGTCTACCTGAACGCTGGACTTGCTGGTGAAC    481
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  AGGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAACGCTGGACTTGCTGGTGGAT    657

Qy  TTGCTGGCCCTGCGCTGCCTGGAAGAATTCAGCCACGCAGCAGCTCTCCTGGG    541
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  TTGCTGGCCCTGCGCTGCCTGGAGAGAATTCTCAGCCACTGCGGCCACTTCCCTGGG    717

Qy  AGTCCGTTTGTGCCAGGTGTCTGGGCTGTTGCCCTGCGTTGCGCTGCCAGGGTCTTCCCTT    601
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  AGCTCCGGCCTCTGCCAGGTGTCTGGGCTGTTGGCCTGCGCCCTGCGCCAGGGTCTTCCCTG    777

Qy  ATCCGCACCCCCCCGGGCTCATCTTAAGGCTGCCCCCTTCCTAACCTACTTTGGA    661
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  ATCCGCACCCCCCCGGGCCATCTCAAGGCTGCCCCCTTCCTCACCTACTTCGGA    837

Qy  TTCAAGTTCACTGAGGGGGCCTTGCTCTCCCAGATTCCTAAACTTCCCTGGCTCC    721
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  TTCCAGGTTCACTGAGGGGGCCTCGGTCTCCCACAGTCGTCCCAGGCTGCGGCTCC    897

Qy  AGCATCACCACCTCCTACCCCCACTCCTCCACCCCCTC-GCTGCTCCTT    780
        |||||||||||||| ||||||||||||||||||||  ||| ||||||
Db  --CCTTCGACAGCTCTCTGGGCACCCGGTCCCCTGCCCACCCTCAGCCGCTCTTT    952

Qy  CCAGTCCTGTCTCTC--TCAAAGGACAGCCAGGAGCTTGTTCACATGTTTCCATTCC-    837
        |||||||||||||    ||||||||||||||||||||||||||||||||||
Db  CCAGACTCGTCCCCTCCTAGAGGCTGCCTGGGCCTGTGCCGTGTGTTTTCCATCCC   1012

Qy  ---ACAGAGGTATCCTTGCTTCTTTAACATCCCATCCACCACAACTATCCCACCTC    891
        |||||||||||| |||||||||||||||||||||||||||||
Db  TAAATACAGTATTCCCACTCTTATCTTACACTCCCCCACCGGCCCACTCTCCCACCTC   1072
```

```
Qy  ACTAGCTCCCAAAGCCCTAC----------------TTATCCTGACTCCCCCACC    892
        |||||||||||||||||            | |||||||||||||||||||
Db  ACTAGCTCCCCCAATCCCTGACCCTTTGAGGCCCCCAGTGATCCTGACTCCCCCCTG   1073

Qy  CACCCGACCACCAGTGTTTATTGACTTTGTGCAC--------------    937
        |||||||||||||||||||||||||||||||
Db  CAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGCCAGGATGGGTCCAGAAG   1133

Qy  ----CAGGCACTGAGATGGGCTGGACCTGGTGGCAGGAAGGCAGAGAACCTGGGA    969
            ||||||||||||||||||||||||||||||||||||||||||||||
Db  CACTTCAGGCACTAAGAGGGGCTGGACCTGGGCGGCAGGAAGCCAAAGGACTGGGC   1193

Qy  GCCGAAGTTCCAACTGTGAGGGGGAAGAGACTGGGACAAGCTCCTCCCTGA--    1024
        ||||||||||||||||||| || ||| || |||||||||||||||||||
Db  GCCAGGAGTTCCCAAATGTGAGGGGCGAGA-AACAAGACAAGCTCCTCCCTTGAGA   1253

Qy  CCTGTGGATTTTGAAA--AGATACTATTTTTATTATTATTGTGACAAAATGTTAA    1080
        |||||||||||||      ||||||||||||||||||||||||||||||||
Db  CCTGTGGATTTTTAAAACAGATATTATTTTTTATTATTATTGTGACAAAATGTTGA   1312
```

RESULT 12
AAV47613
ID   AAV47613 standard; cDNA; 1236 BP.
XX
AC   AAV47613;
XX
DT   27-OCT-1998 (first entry)
XX
DE   TNF related endothelium proliferative agent gene.
XX
KW   ss; TNF; endothelium proliferative agent; TREPA; wound healing; c
KW   tissue grafting; vascularisation; apoptosis; autoimmune; birth cc
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..750
FT   /*tag= a
FT   /product= "TREPA"
XX
PN   WO9835061-A2.
XX
PD   13-AUG-1998.
XX
PF   12-FEB-1998; 98WO-US002859.
XX
PR   12-FEB-1997; 97US-00798692.
PR   10-FEB-1998; 98US-00021706.
XX
PA   (ABBO) ABBOTT LAB.
XX
PI   Wiley SR;
XX
DR   WPI; 1998-447255/38.
DR   P-PSDB; AAW29745.
XX
PT   Detecting nucleic acid encoding TREPA - useful for diagnosis and
PT   treatment of autoimmune disease, tumours and inflammation.
XX
PS   Claim 11; Page 123-4; 142pp; English.
XX
CC   The TNF-related endothelium proliferative agent (TREPA), or its
CC   activators or agonists, are used to treat a deficit of TREPA, e.s
CC   promote wound healing or tissue grafting, by promoting vasculari
CC   also to induce apoptosis for treating cancer and eliminating aut
CC   T cells, as an adjunct to cancer chemotherapy or antiviral treat
CC   TREPA peptides can also be used to target cytotoxic agents or fo
CC   affinity isolation of the corresponding receptor, the nucleic ac
CC   which can be used to transform tumour cells to render them more
CC   responsive to TREPA and to screen for TREPA mimics. Ribozymes, a
CC   RNA, antibodies or peptides, are used to treat TREPA-associated
CC   diseases, e.g. tumours and metastases (by inhibiting vascularisa

r a wide range of autoimmune conditions, conditions
rmal stimulation of epithelial cells (e.g.
s), for birth control (inhibiting ovulation and placental
other angiogenic conditions (e.g. ulcers)

BP; 225 A; 416 C; 358 G; 237 T; 0 U; 0 Other;

51.2%; Score 597.8; DB 2; Length 1236;
rity 75.2%; Pred. No. 2.5e-151;
nservative 0; Mismatches 222; Indels 68; Gaps 8;

```
        GAGCCTGGGCCTGCGCCTGGCCTGCCTTGGCCTCCTGCTGGTGTGTCAGCCTG   61
        |||||||||||||||||||||||||||||||||||||| |||||| |||||||
        GGGCCTGGGCCTGCGCCTGGCCTGCCTCCTGCTGGTTGCCGTGGTCAGTTTG   132

       CTGGGCAACGCTGTCTGCCCAGGAGCCTTCTCAGGAGGAGCTGACACGAGAGAC   121
       ||||||||||||||||||||||||||||||||||||||||| |||||||
       CCGGGCATCGCTGTCGCCCCAGGAGCCTGCCCAGGAGGAGCTGTGTGGGCAGAGGAG   192

      GGAGCCCCCTGAACTGAATCCCCAGACAGAGGAAAGCCAGGATGTGGTACCTTTC   181
      ||||||||||||||||||||||||||||||||||||||||||||||||||
      GGACCCGTCGGAACTGAATCCCCAGACAGAGGAAAGCCAGGATCCTGCGCCTTTC   252

     ACAACTAGTTCCGGCCTGAAGAAGTGCTCTAAAGGCCCGGAAGGCGCCGCCTCGC   241
     |||||||||||||||||||||||  |||||||||||||||||||
     CCGACTAGTTCGGCCTCGAAGAAGTGCACCTAAAGGCCGGAAAACACGGCTCGA   312

    TAATTGCAGCCCATTATGAGGTTCATCCTCGGCCAGGACAGGATGGAGCACAAGCA   301
    |||||| |||||||||||||||||||||||||||||||||||||
    GATCGCAGCCCATTATGAGTTCATCCACGACTTGACAGGACAGGACGAGGCCAGGCA   372

   GGATGGGACAGTGAGTGGCTCGGAAGAGACACCAAAATCAACAGCTTCCAGCCCTCTG   361
   |||||||||||||||||||||||||||||||||||||||||||
   GGACGGGACAGTGAGTGGCTCGGGAGGAAGCCAGATCAACAGCTTCCAGCCCTCTG   432

  CGACCGCCAGATTGGGGAATTTACAGTCATCAGGGCTGGGCTTCTCACTACCTGTAC   421
  ||||||||||||||||||||||||||||||||||||||||||||
  CAACCCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTTCTTACTACCTGTAC   492

 GGTGCACTTTGATGAGGGAAAGCTGTTCACCTGAAGCTGGACTTGCGTGGTGAAC   481
 |||||||||||||||||||||  ||||||||||||||||||||
 GGTGCACTTTGATGAGGGAAGGCTGTCTCACCTGAAGCTGGACTTGCGTGGTGGAT   552

GCTGGCCCTGCGCTGCCTGGAAGAATTCTCAGCCACAGCAGCAAGCTCTTCCTGGG   541
|||||||||||||||||||| ||||||||||||||||||||||
GCTGGCCCTGCGCTGCCTGGAGAATTCTCAGCCACTGCGGCGAGTTCCCTCGGG   612

AGCTCCGTTTGTGCCAGGTGTCTGGGCTTGTCCGCTGCCGCCAGGGTCTTCCCTT   601
|||||||||||||||||||| |||||||||||||||||||||
AGCTCCGCGCCTTCTGCCAGGTGTCTGGGCTGTTGGCCCTGGGCCAGGGTCCTCCCCTG   672

CCGCACCCTCCCCCTGGGGCTCATCTTAAGGCTGCCCCCTCCTAACCTACTTTGGA   661
|||||||||||||||||||||||||||||||||||||||||
CCGCACCCTCCCCCTGGGCCCCATCTCAAGGCTGCCCCCTTCTCTACACTTCCGGA   732

TCAAGTTCACTGAGGGCCTTGCTCTCCCAGATTCCTTCCCAGATTCCTTAAACTTTCCCTGGCTCC   721
|||||||||||||||| ||||||||||||||||||||||||||
CCAGGTTCACTGAGGGGGCCCTGGTCTCCCGCAGTCGTCCCAGCGTGCCCGGCTCC   792

GCATCACCACACCTCCTACCCCACCCCCACTCCTCCACCCCCTC-GCTGCTCCTT   780
||||||||||||| | ||||||||| |||||||
-CCTCGACAGCTCTCTGGACACCCGGTCCCCTTCTGCCCCACCCCTCAGCCGCTCTTT   847

CAGTCCTGTCTCCC--TCAAAGGCAGCCAGACGTTGTTCACATGTTTCCATTCC-   837
|||  |||||| |||| ||||  |||||||||||||||||||||
CAGAGCTTGCCCCCTCCCTCTAGAGGTTGCCTGGCCTGTTTCACGTGTTTTCCATCCC   907

--ACAGACGTATCCTTGCTCTTCTTAACATCCATCCACCACCAACAACTATCCCACCTC   891
|| ||||| ||||||||||||| |||||||||||||||
AAATACAGTATTCCCACTCTTATCTTACACACAACCCCACCCCCACTCTCCACCTC   967

GCTCCCAAAGCCCCTAC---------TTATCCTGACTCCCCACCACT   936
||||| ||| || ||
```

Db 968 ACTAGCTCCCAATCCTGACCCTTTGAGGCCCCAGTGATCTCGACTCCCCCTG(

Qy 937 CACCCGACCACGTGTTATTGACTTGTGCAC-----------------|| |||| ||| |||| |||||

Db 1028 CAGACCCCCAGGGCATTGTTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAG/

Qy 969 -----CAGGCACTGAGATGGGCTGGACCTGGTGCAGGAAGCCAGAGAACCTGGGA(
|| ||||||||||||||||||||||||||||||||||||

Db 1088 CACTTCAGGGCACTAAGAGGGGCTGGACCTGGCCGGCGAGGAAGCCAAAGAGACTGGGC(

Qy 1024 GCCAGAAGTTCCCAACTGTGAGGGGGAAGACTGGGGACAAGCTCCTCCTGGA--
|||| |||| |||||||||||||||||||||||||||||||||||||||

Db 1148 GCCAGGAGTTCCCAAATGTGAGGGGCGAGA-AACAAGACAAGCTCTCCCCTTGAGA

Qy 1080 CCTGTGGATTTGAAAAGATACTATTTT 1108
||||||||||||| |||||||||||

Db 1207 CCTGTGGATTTTTAAAACAGATATATTATT 1235

RESULT 13
AAD04350
ID AAD04350 standard; cDNA; 1236 BP.
XX
AC AAD04350;
XX
DT 04-JUL-2001 (first entry)
XX
DE Human TREPA (TNF related endothelium proliferative agent) cDNA.
XX
KW Human; tumour necrosis factor; TNF; angiogenesis; wound healing;
KW TNF related endothelium proliferative agent; tumour; metastasis;
KW grafting; vulnerary; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..750
FT /*tag= a
FT /product= "Human TREPA (TNF related endothelium
FT proliferative agent)"
XX
PN US6207642-B1.
XX
PD 27-MAR-2001.
XX
PF 26-JUN-1998; 98US-00105343.
XX
PR 12-FEB-1997; 97US-00798692.
PR 10-FEB-1998; 98US-00021706.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Wiley SR;
XX
DR WPI; 2001-280760/29.
DR P-PSDB; AAE00891.
XX
PT Inducing angiogenesis in mammal at desired sites for promoting wo
PT healing, by administering soluble fragment of extracellular domai
PT tumor necrosis factor related endothelium proliferative agent pro
XX
PS Example 2; Col 73-74; 53pp; English.
XX
CC The present invention relates to extracellular signal molecules,
CC particularly members of tumour necrosis factor (TNF) family molec
CC designated as TREPA (TNF related endothelium proliferative agent)
CC Soluble biologically active TREPA are used to treat TREPA-associa
CC diseases, tumours or metastases. TREPA is used for inducing angio
CC in human for promoting wound healing and for vascularising grafte
CC for successful grafting and to promote tissue grafts. The present
CC sequence is a cDNA clone ID #690050 encoding human TREPA
XX
SQ Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 U; 0 Other;

```
        51.2%;  Score 597.8;  DB 4;  Length 1236;
arity   75.2%;  Pred. No. 2.5e-151;
conservative  0;  Mismatches 222;  Indels 68;  Gaps 8;
```

```
Qy   TTGAGCCTGGGCCTGGCGGCTGGCCGTGGCCTGGCCTCCTGCTGGTCGTGGTCAGCCTG   61
         ||||||||||||||||||||||  ||||||||||||||||||  |||||||||||||
Db   TTGGGCCTGGGCCTGGCGTGGCTGGCCTGCCTCCTGGCCTGCCGTGGGTCAGTTTG    132

Qy   GCTGGGCAACGCTGTCTGCCCAGGAGCCTTCTCAGGAGGAGCTGACAGCAGAGGAC    121
         || ||||  ||||||  |||||||||||||||||||  ||| |||||||||||
Db   GCCGGGCATCGCTGTCCGCCCAGGAGCCTGCCCAGGAGGAGCTGTGGTGGCAGAGAG   192

Qy   GGGAGCCCCCTGAACTGAATCCCCAGACAGAGAAAGCCAGGATGTGGTACCTTTC    181
         ||||| |||||||||||||||||   ||||||||||||||||||||||||||
Db   AGGACCCGTCGGAACTGAATCCCCAGACAGAGAAAGCCAGGATCCTGCGGCCTTTC   252

Qy   AAACAACTAGTCCGGCCTCGAAGAAGTGCTCCTAAAGGCCGGAAGGCGCGGCCTCGC    241
         |||||||| ||| |||| |||||| ||| ||||||||||||| |||||| ||||||
Db   AACCGACTAGTTCGGCCTCGAAGAAGTGCACCTAAAGGCCGGAAAACACGGGCTTCGA   312

Qy   GCTATTGCAGCCCATTATGAGGTTCATCCTCGGCCAGGACAGGATGGAGCACAAGCA    301
         || |||||||||| |||| |||||| ||||||||  || ||||||||||| |||
Db   GCGATCGCAGCCCATTATGAAGTTCATCACGACGCCTGGACACGGACGAGCGGCAGGCA   372

Qy   GTGGATGGGACAGTGAGTGGCTGGGAAGAGACCAAAATCAACAGTCCAGCCCTCTG    361
         |||| ||||| |||||||||||||||||||| ||| |||||||||||||||||||
Db   GTGGACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAGCCCTCTG   432

Qy   TACGACCGCCAGATTGGGAATTTACAGTCATCAGGGCTGGGCTCTACTACCTGTAC    421
         ||||||||||||||||||| |||||| |||  |||||||||||||||||||||
Db   TACAACCGCCAGATCCGGGGAGTTTATAGTCACCCGGGGCTGGGCTCTACTACCTGTAC   492

Qy   CAGGTGCACTTTGATGAGGAGAAAGCGTCTACCTGAAGCTGGACTTGCTGGTGAAC    481
         ||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   CAGGTGCACTTTGATGAGGAGAAGGGCTGTCTACCTGAAGCTGGACTTGCTGGTGGAT   552

Qy   GTGCTGGCCCGCTGCGCTGCCTGAAGGAATTCTCAGCCACAGCAGCAAGCTCTCCTGGG    541
         |||||||||| |||||  ||| ||||||| |||||||||  ||||||||||||||||
Db   GTGCTGGCCCCGCTGCCTGCCTGAGAATTCTCAGCCACTGCGGGCGGAGTTCCCTCGGG   612

Qy   CAGTCCGTTTGTGTGCCAGGTGTCTGTGGCTGTTGCCGCTGCGCGCCAGGGTCTTCCCTT    601
         |||||||||||  ||| ||||||||||||||||  ||||||||||||||||||||
Db   CAGCTCCGCCTCTGCCAGGTGTCTGGGCTGTTGGGCTGTGTGGCGCGCCAGGGTCTTCCCTG   672

Qy   ATCCGACCACCTCCCCTGGGCTCATCTTAAGGCTGCCCCCTTCCTAACCTACTTTGGA    661
         |||||||||||||||||||| |||||||||||| |||||||||| |||||||||||
Db   ATCCGACCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCTTCCTCACCTACTTCGGA   732

Qy   TTTCAAGTTCACTGAGGGGCCTTGCTCTCCCAGATTCCTTAAACTTTCCCTGGCTCC    721
         ||||||||||||||||||||| ||| ||||| || || |||||| ||||||||||
Db   TTCCAGGTTCACTGAGGGGCCCTGGTCTCCCCGCAGTCGTCCCAGGCTGCCGGCTCC   792

Qy   AGGCATCACCACACCTCCCTACCCCCACCCCCACTCCTCCACCCCCTC-GCTGCTCCTT    780
         |||  ||||||||  |||||||| ||||| |||||||||||||||||||
Db   --CCTGACAGCTCTCTGGGCACCCGGTCCCGGTCCCCTCTGCGCCCTCGCCGGCTCTTT   847

Qy   CCAGTCCTGTCTCTCC--TCAAAGGCAGCCAGAGCTTGTTCACATGTTCCATTCC-    837
         ||||||   |  ||||     ||||||||||  ||| |||| |||||||||||
Db   CCAGACCTGCCCCTCCCTCTAGAGCTGCCTGGCCTTGTTCACGTGTTTTCCATCCC   907

Qy   ---ACAGACGTATCCTTGCTCTTCTTAACATCCATCCACCACAACTATCCACCTC    891
          ||||||||||||||||||| |||  |||||||||||||||||||||
Db   TAAATACAGTATTCCCACTCTTATCTTACAACAACCCACCGCCCACTCTCCACCTC   967

Qy   AGCTCCCAAAGCCCCTAC--------TTATCCCTGACTCCCCACCCACT    936
          ||||  ||||  |||       |||||||||||||||||||
Db   AGCTCCCCAATCCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCCCCCCTGGCCA   1027

Qy   CCGACCACGTGTTTATTGACTTTGTGCAC---------    968
          ||||||  ||| |||||| ||||
Db   AACCCCAGGCAATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGGTCCAGAGACCC   1087
```

```
Qy   969  -----CAGGCACTGAGATGGCTGGGACCTGGTGGCAGGAAGCCAGAGAACCTGGG(
          |||  |||| ||||  ||||||||||||||||||||||||||||||||||
Db   1088 CACTTCAGGCACTAAGAGGGGCTGGGACCTGGCGGGCAGGAAGCCAAAGAGACTGGG(

Qy   1024 GCCAGAAGTTCCCAACTGTGAGGGGGGAAGAGTGGGGACAAGCTCCTCCCTTGGA--
          ||||||||| |||   ||| ||| |||| ||||||||  ||| ||||||||||||
Db   1148 GCCAGGAGTTCCCAAATGTGAGGGCGCGAGA-AACAAGACAAGCTCCTCCCTTGAG/

Qy   1080 CCTGTGGATTTGAAAAGATACTATTTT 1108
          |||||||||||    ||||| |||||||
Db   1207 CCTGTGGGATTTTTAAAACAGATATATT 1235
```

RESULT 14
AAX23424
ID   AAX23424 standard; DNA; 1030 BP.
XX
AC   AAX23424;
XX
DT   18-JUN-1999 (first entry)
XX
DE   Human TNRL3 DNA.
XX
KW   Tumour necrosis factor receptor; signal transducer molecule; TNF
KW   developmental abnormality; gestational abnormalitity; prostate c.
KW   APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy;
KW   cytoplasmic domain; immunogen; antibody preparation; breast carc
KW   apoptosis; human; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..627
FT   /*tag= a
FT   /product= "TNRL3"
XX
PN   WO911791-A2.
XX
PD   11-MAR-1999.
XX
PF   04-SEP-1998; 98WO-US018393.
XX
PR   05-SEP-1997; 97US-00924634.
XX
PA   (UNIW) UNIV WASHINGTON.
XX
PI   Chaudhary PM;
XX
DR   WPI; 1999-205191/17.
DR   P-PSDB; AAW93590.
XX
PT   New Tumor Necrosis Factor family receptor polypeptides and ligan
PT   useful for diagnosis and treatment of prostate cancer and develo
PT   or gestational abnormalities.
XX
PS   Example VII; Fig 13A; 156pp; English.
XX
CC   This invention describes isolated Tumor Necrosis Factor (TNF) fa
CC   receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC   fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and T
CC   their active fragments. APO4 is useful for diagnosing prostate c
CC   determining levels of APO4 in an individual. Prostate cancer can
CC   treated using APO4 selective binding agents linked to a therapeu
CC   moiety. APO4 polypeptides are also useful for identifying select
CC   binding agents, useful in diagnosis/treatment of disease by bind
CC   agents to the polypeptide/active fragment which is extracellular
CC   expressed on the cell surface. The binding is preferably perform
CC   vivo. APO4 polypeptides/ active fragments are also useful for sc
CC   for agonists and antagonists by binding and observing the change
CC   activity. Effective pharmacological agents useful in diagnosis o
CC   treatment of disease are also identified using APO4 polypeptides
CC   fragments and APO4 signal transducer molecules that specifically
CC   with a cytoplasmic domain of APO4 and detecting a change in leve

method is performed in vivo or in vitro. APO polypeptides
as immunogens for preparing antibodies. APO4 is also
gnosis/treatment of developmental or gestational
APO8 was transfected to human breast carcinoma cell line
ced apoptosis

BP; 223 A; 317 C; 279 G; 211 T; 0 U; 0 Other;

44.8%;   Score 522.8; DB 2; Length 1030;
rity 81.6%; Pred.No.4.8e-131;
nservative 0; Mismatches 137; Indels 8; Gaps 3;

CCTGGGGAGCTGGCAACGCTGTCTGCCCAGGAGCCTTCTCAGGAGGAGCTGACA 112
||||| ||||||||| ||||||||||||||||||||||||||||||||||||||
TTTGGGGAGCTGGCGCATCGCTGTCCGCCCCAGGAGCCTGCCCCAGGAGGAGCTGGTG 60

GGACCGCCGGGAGCCCCCTGAACTGAATCCCCAGACAGAGGAAAGCCAGGATGTG 172
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGAGGACCAGGACCCGTCGGGAACTGAATCCCCAGACACAGAAGAGGAAGCCAGGATCCT 120

TTTCTTGGAACAACTAGTCCGGCCTGGAAGAAGTGCTCCTAAAGGCCGAAGGCG 232
||||| ||| |||||||||||||||||| |||||| |||||||||||||||||
TTTCCTGAACCGACTAGTTCGGCCTCGCAGAGTGCACCTAAAGGCCGGAAAACA 180

TCCCCGAGCTATTGCAGCCCATTATGAGGTTCATCCTCGGCCAGGACAGGATGA 292
||||||||||||||||||||||||||||||| |||||||||||||||||||||
TCGAAGAGCGATCGCAGCCCATTATGAGTTCATCCACGACCTGGACACGGACGA 240

AGCAGGTGTGGATGGGACAGTGAGTGGCTGGGAAGAGACCAAAATCAACAGCTCC 352
|||||||||||||||||||||||||||||||||||||||||||||||||||||
GGCAGGTGTGGACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAAATCAACAGCTCC 300

TCTGCGCTACGACGCCAGATTGGGGAATTTACAGTCATCATCAGGGCTGGCTCTAC 412
||||||||||||||||||||||||||| |||| |||||||||||||||||||||
TCTGCGCTACTACCGCCAGATCGGGGAGTTTATAGTCACCCGGGGCTGGCTCTAC 360

GTACTGTCAGGTGCACTTTGATGAGGGAAAGGCTGTCTACCTGAAGTGGACTTG 472
||||||||||||||||||||||||||| |||||||||||||||||||||||||
GTACTGTCAGGTGCACTTTGATGAGGGGGAGGCTGTCTACCTGAAGTGGACTTG 420

GAACGGTGTGCTGCCCCTGCGCTGCCTGGAAGAATTCTCAGCCACAGCAGCAAGC 532
|||||||||||||||||||||||||||||||||||||||||||||||||||||
GGATGGTGTGCTGCCCCTGCGCTGCCTGGAGGAATTCTCAGCCCACTGCGCCCAGT 480

TGGGCCCCAGCTCCGTTCTGCCAGGTGCTGCTGGGCGTGTTGCCGCTGCCGCTGCCAGGG 592
|||||||||||||||||||||||||||||||||||||||||||| ||||||||||
CGGGCCCCAGCTCCGCCCTGCCAGGTGCTGGGGCTGTTGGCCCTGCCGCCGCCAGGG 540

CCTTCGGATCCGCACCCTCCCCCTGGGCTCATCTTAAGGCTGCGCCCCTTCCTAACC 652
|||||||||||||||||||||||||||||||||||||||||||||||||||||
CCTGCGGATCCGCACCCTCCCCCTGGGCCCATCTCCAAGGCTGCCCCCTTCCTCACC 600

TGGACTCTTTCAAGTCACTGAGGGGCCTTGCTCTCCCAGATTCCTTAAACTTTC 712
||||||||||||||||||||||||||||||||||||||||||||||||||||
CGGACTCTTCCAGTTCACTGAGGTTCACTGAGGGCCCTGGTCTCCCACAGTCGTCCCAGGCT 660

GCTCCAGGAGCATCACCACACTCCCTACCCACCCCCACTCCTCCACCCCCTC-G 771
|||||||| ||||||||||||||||||||||||||||||||||||||||||
GCTCC---CCTCGACAGCTCTCTGGGCACCCGGTCCCCTCTGCCCCCACCCTCAG 715

CCTTGGTCCAGTCCTGTCTCTCC--TCAAAGGCAGCCAGAGCTTGTTCACATGTT 829
|||||| |||| |||||| |||||||||||||||||||||||||||||||
TCTTTGCTCCAGACCTGCCCTCCCCTAGAGGCTGCCTGGGCCTGGCCTGTTCACGTGTT 775

TTCC 837

ATCC 783

dard; DNA; 898 BP.

```
                42.7%;    Score 498.8;  DB 4;   Length 898;
arity          87.0%;    Pred. No. 1.4e-124;
onservative       0;    Mismatches  82;    Indels     0;   Gaps     0;

TGGGGAGCTGGGCAACGCTGTCTGCCCAGGAGCCTTCTCAGGAGGAGCTGACAGCA 115
|||||||| ||||||| ||||||||||||||||||| ||||||||||| |||||||
TGGGGAGCCGGGGATCGCTGTCCGCCCAGGAGCCTGCCCAGGAGGAGCTGGTGGCA 309

ACCCGCGGGAGCCCCCTGAACTGAATCCCAGACAGAGGAAAGCCAGGATGTGGTA 175
||||||| ||||||||||||||||||||||||||||||||||| |
AAGGACCAGGACCCGTCGGAACTGAATCCCAGACAGAAGCCAGGATCCTGCG 369

TTCTTGGACAACTAGTCCGGCCTCGAAGAAGTGCTTCCTAAAGGCCGGAAGGCGCGG 235
|||| ||||||||||||||||||||||||||||||||||| |||||||||||
TCCTGAACCGACTAGTTCGGCCCTCGCAGAAGTGCACCTAAAGGCCGGAAAAACACGG 429

TGCCGAGCTATTGCAGCCCATTATGAGGTTCATCCTCGGCCAGGACAGGATGGAGCA 295
||| | |||||||||||||||||||||||||||| || || ||||||| ||||
TGAAGAGGCGGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACGGAGCG 489

TCAGGTGTGGATGGGACAGTGAGTGGCTGGGAAGAGACCAAAATCAACAGCTCCAGC 355
||||||||| |||||||||||||||||||||||| ||| ||||||||||||||||
GCAGGTGTGACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAGC 549

TTGCGCTACGACCGCCAGATTGGGGAATTACAGTCATCAGGGCTGGGCTCTACTAC 415
||||||||| |||||||||||||||| |||||| |||||||||||||||||||||
TTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTACTAC 609

ACTGTCAGGTGCACTTGATGAGGGAAAGGCTGTCTACCTGAAGGCTGGACTTGCTG 475
|||||||||||||||||||||||||||||||||||||||| ||||||||||||||
ACTGTCAGGTGCACTTGATGAGGGAAAGGCTGTCTACCTGAAGGCTGGACTTGCTG 669

ACGGTGTGTGCTGGCCCTGCCGCTGCCTGGAGAATTCTCAGCCACAGCAGCAAGCTCT 535
|||||||||||||||| |||||||||||||||||||| ||| || |||||||||
GATGGTGTGCTGGCCCCGCCTGCCGCTGCCTGGAGAATTCTCAGCCACTGCGGCCAGTTCC 729

GGGCCCCAGCTCCGTTTGTGCCAGGTGTCTGGGCTGTTGGCCGTCGGGGCCAGGGTCT 595
||||||||||||||||||||||| |||||||||||||||||||||||||||||||
GGGCCCCAGCTCCGCGCTCTGCCAGGTGTCTGGGCTGTTGGCCCTGCGGGCCAGGGTCC 789

TTTCGGATCGCACCCTCCCCTGGGCTCATCTTAAGGCTGCCCCTTCCTAACCTAC 655
||||||||||||||||||||||||||||||||||||||||| |||||||
TTGCCGGATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCTTCCTCACCTAC 849

GGACTCTTTCAAGTTCACTGAGGGGCC 685
||||| ||||||||||||||||||
GGACTCTTTCCAGGTTCACTGAGCGCGCC 879
```

April 7, 2004, 21:32:17
1 secs

```
        GenCore version 5.1.6
opyright (c) 1993 - 2004  Compugen Ltd.

n search, using sw model

il 7, 2004, 17:38:07 ; Search time 9.28291 Seconds
                       (without alignments)
                       1262.081 Million cell updates/sec

09-245-198A-2
2
TLSLGLALACLGLLLVVVSL..........PWAHLKAAPFLTYFGLFQVH 225

SUM62
op 10.0 , Gapext 0.5

681 seqs, 52070155 residues

s satisfying chosen parameters:     141681

th: 0
th: 2000000000

nimum Match 0%
ximum Match 100%
ting first 45 summaries

issProt_42:*

the number of results predicted by chance to have a
 than or equal to the score of the result being printed,
ed by analysis of the total score distribution.

                SUMMARIES
```

ery	ch Length	DB	ID	Description
0.0	225	1	TN12_MOUSE	O54907 mus musculu
0.8	249	1	TN12_HUMAN	O43508 homo sapien
9.3	272	1	TNF5_CHICK	Q9i8d8 gallus gall
3.8	316	1	TN11_MOUSE	O35235 m tumor nec
7.8	260	1	TNF5_CANFA	O97626 canis famil
7.8	318	1	TN11_RAT	Q9ese2 r tumor nec
7.6	532	1	PPBN_HUMAN	P10696 homo sapien
7.6	3664	1	MINT_HUMAN	Q96t58 homo sapien
7.5	244	1	TNFC_HUMAN	Q06643 homo sapien
7.5	244	1	TNFC_PANTR	Q862z7 pan troglod
7.5	261	1	TNF5_AOTTR	Q9bdm3 aotus trivi
7.5	261	1	TNF5_CALJA	Q9bdn3 callithrix
7.4	240	1	TN14_HUMAN	O43557 homo sapien
7.4	1237	1	B3A2_MOUSE	P13808 mus musculu
7.4	261	1	TNF5_HUMAN	P29965 homo sapien
7.4	261	1	TNF5_MACMU	Q9bdc7 macaca mula
7.3	240	1	TNF5_MACNE	Q9bdm7 macaca neme
7.3	1237	1	B3A2_RABIT	P48746 oryctolagus
7.2	278	1	TNF6_RAT	P36940 rattus norv
7.2	241	1	TN13_MOUSE	Q9d777 mus musculu
7.2	250	1	TNFC_MACEU	Q9xt47 macropus eu
7.1	400	1	TRPB_CHRVO	Q7nud8 chromobacte
7.1	2779	1	LVA_DROME	Q8ms1 drosophila
7.1	535	1	PPBI_HUMAN	P05187 homo sapien
6.9	246	1	C1QC_MOUSE	Q02105 mus musculu
6.9	1234	1	B3A2_RAT	P23347 rattus norv
6.8	1465	1	DPO3_STRMU	Q8dwe0 streptococc
6.8	197	1	TNFB_RABIT	P10154 oryctolagus
6.8	300	1	NTH1_MOUSE	O35980 mus musculu
6.8	920	1	PARC_SYNY3	P73077 synechocyst
6.8	1584	1	U104_CAEEL	P23678 caenorhabdi
6.8	2468	1	MARPB_HUMAN	P46821 homo sapien
6.8	260	1	TNF5_FELCA	O97605 felis silve

```
34  78.5  6.8   480  1  KCG2_RAT      Q9qyu3 ratt
35  78.5  6.8  1164  1  PHYD_ARATH    P42497 arab
36  78    6.7   285  1  T13B_HUMAN    Q9y275 homo
37  78    6.7   495  1  GATB_METAC    Q8thj0 meth
38  78    6.7   495  1  GATB_METMA    Q8pxj0 meth
39  78    6.7   763  1  APP2_HUMAN    Q06481 homo
40  77.5  6.7   201  1  TNFB_MACEU    Q9xt48 macr
41  77.5  6.7   261  1  TNF5_PIG      Q95mq5 sus
42  77    6.6   279  1  TNF6_MOUSE    P41047 mus
43  77    6.6   817  1  NAH1_BOVIN    Q28036 bos
44  76.5  6.6   214  1  SMP_ECOLI     P18838 esch
45  76.5  6.6   788  1  NASP_HUMAN    P49321 homo
```

ALIGNMENTS

```
RESULT 1
TN12_MOUSE
ID   TN12_MOUSE     STANDARD;     PRT;     225 AA.
AC   O54907; Q9CTP2;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Tumor necrosis factor ligand superfamily member 12 (TNF-related w
DE   inducer of apoptosis) (TWEAK) (Fragment).
GN   TNFSF12.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Peritoneal macrophage;
RX   MEDLINE=98070415; PubMed=9405449;
RA   Chicheportiche Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H.,
RA   Hession C., Garcia I., Browning J.L.;
RT   "TWEAK, a new secreted ligand in the tumor necrosis factor family
RT   weakly induces apoptosis.";
RL   J. Biol. Chem. 272:32401-32410(1997).
RN   [2]
RP   SEQUENCE OF 83-225 FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Retina;
RX   MEDLINE=21085660; PubMed=1217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kasukawa T., Saito R.
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washic
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto K.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilmir
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
CC   -!- FUNCTION: Binds to FN14 and possibly also to TNRFSF12/APO3.
CC       inducer of apoptosis in some cell types. Promotes angiogenesi
CC       the proliferation of endothelial cells. Mediates NF-KappaB
CC       activation (By similarity).
CC   -!- SUBUNIT: Homotrimer (Potential).
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein and secreted
CC       similarity).
CC   -!- TISSUE SPECIFICITY: Widely expressed.
CC   -!- PTM: The soluble form is produced from the membrane form by
```

ic processing (By similarity).
Y: Belongs to the tumor necrosis factor family.

0; AAC53517.1; -.
9; BAB32249.1; -.
259; Tnfsfl2.
006052; TNF_family.
008983; TNF_like.
); TNF; 1.
07; TNF; 1.
251; TNF_1; FALSE_NEG.
049; TNF_2; 1.
iogenesis; Apoptosis; Transmembrane; Glycoprotein;

```
1    1      TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
            MEMBER 12, MEMBRANE FORM.
<1   225    TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
70   225    MEMBER 12, SECRETED FORM (BY SIMILARITY).
<1   21     SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
            (POTENTIAL).
22   225    EXTRACELLULAR (POTENTIAL).
69   70     CLEAVAGE (BY SIMILARITY).
67   186    POTENTIAL.
15   115    N-LINKED (GLCNAC...) (POTENTIAL).
25 AA;  90C412CC0480659B CRC64;
```

```
arity      100.0%;   Score 1162;  DB 1;  Length 225;
           100.0%;   Pred. No. 2.5e-95;
onservative  0;   Mismatches  0;   Indels  0;   Gaps  0;

GLALACLGLLLVVVSLGSWATLSAQEPSQEELTAEDRREPPELNPQTEESQDVVPF 60
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
GLALACLGLLLVVVSLGSWATLSAQEPSQEELTAEDRREPPELNPQTEESQDVVPF 60

VRPRRSAPKGRKARPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSSSPL 120
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
VRPRRSAPKGRKARPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSSSPL 120

RQIGEFTVIRAGLYYLYCQVHFDEGKAVYLKDLLVNGVLALRCLEEFSATAASSPG 180
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
RQIGEFTVIRAGLYYLYCQVHFDEGKAVYLKDLLVNGVLALRCLEEFSATAASSPG 180

RLCQVSGLLPLRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 225
||||||||||||||||||||||||||||||||||||||||||
RLCQVSGLLPLRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 225
```

Z7;   STANDARD;   PRT;   249 AA.

(Rel. 41, Created)
(Rel. 41, Last sequence update)
(Rel. 42, Last annotation update)
is factor ligand superfamily member 12 (TNF-related weak
poptosis) (TWEAK) (APO3 ligand).
PO3L OR DR3LG.
(Human).
etazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
theria; Primates; Catarrhini; Hominidae; Homo.
606;

M N.A., AND N-TERMINUS OF SOLUBLE FORM.
liver, and Tonsil;

RX   MEDLINE=98070415; PubMed=9405449;
RA   Chicheportiche Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H.,
RA   Hession C., Garcia I., Browning J.L.;
RT   "TWEAK, a new secreted ligand in the tumor necrosis factor famil
RT   weakly induces apoptosis.";
RL   J. Biol. Chem. 272:32401-32410(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal kidney;
RX   MEDLINE=98228355; PubMed=9560343;
RA   Marsters S.A., Sheridan J.P., Pitti R.M., Brush J., Goddard A.,
RA   Ashkenazi A.;
RT   "Identification of a ligand for the death-domain-containing rece
RT   Apo3.";
RL   Curr. Biol. 8:525-528(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Tonsil;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanche
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [4]
RP   FUNCTION.
RX   MEDLINE=99185061; PubMed=10085077;
RA   Lynch C.N., Wang Y.C., Lund J.K., Chen Y.-W., Leal J.A., Wiley S
RT   "TWEAK induces angiogenesis and proliferation of endothelial cel
RL   J. Biol. Chem. 274:8455-8459(1999).

CC   -!- FUNCTION: Binds to FN14 and possibly also to TNRFSF12/APO3.
CC   inducer of apoptosis in some cell types. Mediates NF-KappaB
CC   activation. May promote angiogenesis and the proliferation o
CC   endothelial cells.
CC   -!- SUBUNIT: Homotrimer (Potential).
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein and secreted.
CC   -!- TISSUE SPECIFICITY: Highly expressed in adult heart, pancrea
CC   skeletal muscle, brain, colon, small intestine, lung, ovary,
CC   prostate, spleen, lymph node, appendix and peripheral blood
CC   lymphocytes. Low expression in kidney, testis, liver, placen
CC   thymus and bone marrow. Also detected in fetal kidney, liver
CC   lung and brain.
CC   -!- PTM: The soluble form derives from the membrane form
CC   by proteolytic processing.
CC   -!- SIMILARITY: Belongs to the tumor necrosis factor family.
CC   -!- CAUTION: Ref.3 sequence differs from that shown due to a
CC   frameshift in position 125.

DR   EMBL; AF030099; AAC51923.1; -.
DR   EMBL; AF055872; AAC39724.1; -.

```
CC   -!- SUBUNIT: Homotrimer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein. Also exists a
CC       extracellular soluble form (By similarity).
CC   -!- PTM: The soluble form derives from the membrane form by
CC       proteolytic processing (By similarity).
CC   -!- SIMILARITY: Belongs to the tumor necrosis factor family.
CC
CC   This SWISS-PROT entry is copyright. It is produced through a coll
CC   between the Swiss Institute of Bioinformatics and the EMBL out
CC   the European Bioinformatics Institute. There are no restriction
CC   use by non-profit institutions as long as its content is in
CC   modified and this statement is not removed. Usage by and for c
CC   entities requires a license agreement (See http://www.isb-sib.ch/
CC   or send an email to license@isb-sib.ch).
CC
DR   EMBL; AJ243435; CAB95748.2; -.
DR   HSSP; P29965; 1ALY.
DR   GO; GO:0016021; C:integral to membrane; ISS.
DR   GO; GO:0005174; F:CD40 receptor binding; ISS.
DR   GO; GO:0042100; P:B-cell proliferation; ISS.
DR   GO; GO:0006954; P:inflammatory response; ISS.
DR   GO; GO:0007159; P:leukocyte cell adhesion; ISS.
DR   GO; GO:0030168; P:platelet activation; ISS.
DR   InterPro; IPR003263; TNF_5.
DR   InterPro; IPR006052; TNF_family.
DR   InterPro; IPR008983; TNF_like.
DR   InterPro; IPR003636; TNF_subf.
DR   Pfam; PF00229; TNF; 1.
DR   PRINTS; PR01702; CD40LIGAND.
DR   ProDom; PD008600; TNF_5; 1.
DR   ProDom; PD002012; TNF_subf; 1.
DR   SMART; SM00207; TNF; 1.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 1.
KW   Cytokine; Transmembrane; Glycoprotein; Signal-anchor.
FT   CHAIN        1    272       TUMOR NECROSIS FACTOR LIGAND SUPERFA
FT                               MEMBER 5, MEMBRANE FORM.
FT   CHAIN     -111    272       TUMOR NECROSIS FACTOR LIGAND SUPERFA
FT                               MEMBER 5, SOLUBLE FORM (BY SIMILARIT
FT   DOMAIN       1     23       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    24     44       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROT
FT                               (POTENTIAL).
FT   DOMAIN      45    272       EXTRACELLULAR (POTENTIAL).
FT   SITE       110    111       CLEAVAGE (BY SIMILARITY).
FT   DISULFID   190    229       POTENTIAL.
FT   CARBOHYD   124    124       N-LINKED (GLCNAC. . .)   (POTENTIAL).
FT   CARBOHYD   146    146       N-LINKED (GLCNAC. . .)   (POTENTIAL).
FT   CARBOHYD   251    251       N-LINKED (GLCNAC. . .)   (POTENTIAL).
SQ   SEQUENCE   272 AA;   30832 MW;  8CD0338A924E044B CRC64;
```

Query Match   9.3%;   Score 108.5; DB 1; Length 272;
Best Local Similarity 22.5%;   Pred. No. 0.024;;
Matches   58; Conservative   40; Mismatches 107; Indels   53; C

Qy   1 VLSLGLALACLGLLLVVVSLGSWATLSAQ---EPSQEELTAEDRREP-------
Db   34 VQTIGTVLFCLYLHMKMDKMEEVLSLNEDYIFLRKVQKCQTGEDQKSTLLDCEKVI

Qy   46 ELNPQTEESQDVVPFLEQLVRPRRSAPKGRK----ARPRRAIAAHYEVHPRPGQI
Db   94 DLQCKDRTASEELPKFEMHRGHEHPHLKSRNETSVAEEKRQPIATHLA-------

Qy   101 GVDGTVSGWEETK-INSSSPLRYDRQIGEFTVIRAGLYYLYCQVHFDEGKA-----
Db   146 NTTVRVLKWMTSYAPTSSLISYHE--GKLKVBKAGLYYIYSQVSFCTKAAASAPI

Qy   152 YLKLDLLVNGVLALRCLEEFSATAASSPGPQLRLCQV-----SGLLLRPGSSLR
Db   204 YLYLPMEEDRLL-MKGLDTHSTSTA------LCEQSIREGGVFELRQGDMVF

Qy   207 WAHLKAAPFLTYFGLFQV 224
Db   255 STAVNVNPGNTYFGMFKL 272

STANDARD;   PRT;   316 AA.
06; Q9JUK8; Q9JUK9; Q9R1Y0;
(Rel. 40, Created)
(Rel. 40, Last sequence update)
(Rel. 42, Last annotation update)
is factor ligand superfamily member 11 (Receptor activator
actor kappa B ligand) (RANKL) (TNF-related activation-
cine) (TRANCE) (Osteoprotegerin ligand) (OPGL) (Osteoclast
ion factor) (ODF) (Osteoclastogenesis-inhibitory factor)

ANKL OR TRANCE OR OPGL.
(Mouse).
etazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
theria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
090;

M N.A. (ISOFORM 1).
loma;
0112; PubMed=9312132;
no J., Arron J., Robinson E., Orlinick J., Chao M.,
, Cayani E., Bartlett F.S. III, Frankel W.N., Lee S.Y.,

novel ligand of the tumor necrosis factor receptor family
as c-Jun N-terminal kinase in T cells.";
n. 272:25190-25194(1997).

M N.A. (ISOFORM 1).
c lymphoma;
2977; PubMed=9367155;
, Maraskovsky E., Billingsley W.L., Dougall W.C.,
, Roux E.R., Teepe M.C., DuBose R.F., Cosman D.,

of the TNF receptor and its ligand enhance T-cell growth
c-cell function.";
75-179(1997).

M N.A. (ISOFORM 1).
marrow;
7661; PubMed=9568710;
Timms E., Tan H.-L., Kelley M.J., Dunstan C.R.,
Elliott R., Colombero A., Elliott G., Scully S., Hsu H.,
Hawkins N., Davy E., Capparelli C., Eli A., Qian Y.-X.,
Sarosi I., Shalhoub V., Senaldi G., Guo J., Delaney J.,

erin ligand is a cytokine that regulates osteoclast
ion and activation.";
176(1998).

M N.A. (ISOFORM 1).
marrow stroma;
8248; PubMed=9520411;
hima N., Nakagawa N., Yamaguchi K., Kinosaki M.,
-I., Tomoyasu A., Yano K., Goto M., Murakami A., Tsuda E.,
Higashio K., Udagawa N., Takahashi N., Suda T.;
differentiation factor is a ligand for
rin/osteoclastogenesis-inhibitory factor and is identical
NKL.";
Acad. Sci. U.S.A. 95:3597-3602(1998).

M N.A. (ISOFORM 1).
4075; PubMed=10196481;
Kodaira K., Mizuno A., Yasuda H., Shima N., Murakami A.,
ashio K.;
characterization of the gene encoding mouse osteoclast
ion factor.";
-127(1999).

M N.A. (ISOFORMS 1; 2 AND 3).

RX   MEDLINE=21150053; PubMed=11250921;
RA   Ikeda T., Kasai M., Utsuyama M., Hirokawa K.;
RT   "Determination of three isoforms of the receptor activator of nu
RT   factor-kappaB ligand and their differential expression in bone a
RT   thymus.";
RL   Endocrinology 142:1419-1426(2001).
RN   [7]
RP   SEQUENCE OF 139-147, PROCESSING, AND N-GLYCOSYLATION.
RX   MEDLINE=9240759; PubMed=10224132;
RA   Lum L., Wong B.R., Josien R., Becherer J.D., Erdjument-Bromage H
RA   Schloendorff J., Tempst P., Choi Y., Blobel C.P.;
RT   "Evidence for a role of a tumor necrosis factor-alpha
RT   (TNF-alpha)-converting enzyme-like protease in shedding of TRANC
RT   TNF family member involved in osteoclastogenesis and dendritic c
RT   survival.";
RL   J. Biol. Chem. 274:13613-13618(1999).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 158-316.
RX   MEDLINE=21464816; PubMed=11581298;
RA   Lam J., Nelson C.A., Ross F.P., Teitelbaum S.L., Fremont D.H.;
RT   "Crystal structure of the TRANCE/RANKL cytokine reveals determin
RT   of receptor-ligand specificity.";
RL   J. Clin. Invest. 108:971-979(2001).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 137-316.
RX   MEDLINE=21839021; PubMed=11733492;
RA   Ito S., Wakabayashi K., Ubukata O., Hayashi S., Okada F., Hata T
RT   "Crystal structure of the extracellular domain of mouse RANK lig
RT   2.2-A resolution.";
RL   J. Biol. Chem. 277:6631-6636(2002).
CC   -!- FUNCTION: Cytokine that binds to TNFRSF11B/OPG and to
CC   TNFRSF11A/RANK. Osteoclast differentiation and activation fa
CC   Augments the ability of dendritic cells to stimulate naive T
CC   proliferation. May be an important regulator of interactions
CC   between T cells and dendritic cells and may play a role in t
CC   regulation of the T cell-dependent immune response. May also
CC   an important role in enhanced bone-resorption in humoral
CC   hypercalcemia of malignancy.
CC   -!- SUBUNIT: Homotrimer.
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein and secreted
CC   (isoforms 1 and 2); Cytoplasmic (isoform 3).
CC   -!- ALTERNATIVE PRODUCTS:
CC   Event=Alternative splicing; Named isoforms=3;
CC   Name=1;
CC   IsoId=O35235-1; Sequence=Displayed;
CC   Name=2;
CC   IsoId=O35235-2; Sequence=VSP_006449;
CC   Name=3;
CC   IsoId=O35235-3; Sequence=VSP_006448;
CC   -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THYMUS AND LYMPH NOD
CC   NOT IN NONLYMPHOID TISSUES AND IS ABUNDANTLY EXPRESSED IN T
CC   BUT NOT IN B CELLS. A HIGH LEVEL EXPRESSION IS ALSO SEEN IN
CC   TRABECULAR BONE AND LUNG.
CC   -!- PTM: N-glycosylated.
CC   -!- PTM: The soluble form of isoform 1 derives from the membrane
CC   by proteolytic processing. The cleavage may be catalyzed by
CC   ADAM17. A further shorter soluble form was observed.
CC   -!- DISEASE: Deficiency in TNFSF11 results in failure to form lc
CC   alveolar mammary structures during pregnancy, resulting in d
CC   of newborns. Trance-deficient mice show severe osteopetrosis
CC   no osteoclasts, marrow spaces, or tooth eruption, and exhibi
CC   profound growth retardation at several skeletal sites, inclu
CC   the limbs, skull, and vertebrae and have marked chondrodyspl
CC   with thick, irregular growth plates and a relative increase
CC   hypertrophic chondrocytes.
CC   -!- SIMILARITY: Belongs to the tumor necrosis factor family.

il to license@isb-sib.ch).

```
Db    AAC71061.1; -.
;     AAB86812.1; -.
;     AAC40113.1; -.
;     BAA25425.1; -.
;     BAA36970.1; -.
;     BAA36970.1; JOINED.
;     BAA36970.1; JOINED.
;     BAA36970.1; JOINED.
;     BAA97257.1; -.
;     BAA97258.1; -.
;     BAA97259.1; -.
-JAN-03.
089; Tnfsfl1.
3; P:organogenesis; IMP.
7; P:ossification; IMP.
006052; TNF family.
008983; TNF_like.
003636; TNF_subf.
; TNF; 1.
012; TNF_subf; 1.
7; TNF; 1.
251; TNF_1; FALSE_NEG.
049; TNF_2; 1.
ferentiation; Receptor; Glycoprotein; Transmembrane;
; 3D-structure; Alternative splicing.
1     316              TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
                       MEMBER 11, MEMBRANE FORM.
39    316              TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
                       MEMBER 11, SOLUBLE FORM.
1     48               CYTOPLASMIC (POTENTIAL).
49    69               SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
                       (POTENTIAL).
70    316              EXTRACELLULAR (POTENTIAL).
38    139              CLEAVAGE.
97    197              N-LINKED (GLCNAC. . .)    (POTENTIAL).
62    262              N-LINKED (GLCNAC. . .)    (POTENTIAL).
1     117              Missing (in isoform 3).
                       /FTId=VSP 006448.
14    44               SEEMGSGPGVPHEGPLHPAPSAPAPAPPA -> TP (in
                       isoform 2).
                       /FTId=VSP 006449.
99    99               G -> D (IN REF. 2).
41    143              MISSING (IN REF. 5).
64    169
71    172
81    182
86    187
91    192
94    196
98    201
02    203
02    207
11    224
25    227
34    245
```

larity   8.8%;   Score 102.5;   DB 1;   Length 316;
onservative   24.3%;   Pred. No. 0.096;   Mismatches 107;   Indels   65;   Gaps   13;

```
                                                                        -----REPPELNPQT 51
-ALACLGLLLVVSLGSWATLSAQ-EPSQELTAEDR-
  |||||| || ||| || :||: || |:|: |||
LALLGLGLGQVVCSIALFLYFRAQMDPNR---ISEDSTHCFYRLRLHENAGLQDST 104

QDVVP----------FLEQLVRPRR--SAPK--------GRKARPRRAIA 84
  |:|:||                                  :  | || ||
EDTLPDSCRRMKQAFQGAVQKELQHIVGPQRPSGAPAMMEGSWLDVAQRGKPEAQPF 164

EVHPRPGQDGAQAGVDGTVSGWEE----TKINSSSPLRYDRQIGEFTVIRAGLYYLY 140
  :| : ||||        |||| | || :  | ||||||
```

RESULT 5
TNF5_CANFA
```
ID    TNF5_CANFA    STANDARD;    PRT;    260 AA.
AC    O97626;
DT    16-OCT-2001 (Rel. 40, Created)
DT    16-OCT-2001 (Rel. 40, Last sequence update)
DT    28-FEB-2003 (Rel. 41, Last annotation update)
DE    Tumor necrosis factor ligand superfamily member 5 (CD40 ligand).
GN    TNFSF5 OR CD40LG OR CD40L.
OS    Canis familiaris (Dog).
OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC    Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX    NCBI_TaxID=9615;
RN    [1]
RP    SEQUENCE FROM N.A.
RA    Hosie M.H., Willett B.J.;
RT    "Adjuvant properties of canine CD40L.";
RL    Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
CC    -!- FUNCTION: Cytokine that binds to TNFRSF5. Mediates B-cell
CC        proliferation in the absence of co-stimulus as well as IgE
CC        production in the presence of IL-4. Involved in immunoglobuli
CC        class switching (By similarity).
CC    -!- SUBUNIT: Homotrimer (By similarity).
CC    -!- SUBCELLULAR LOCATION: Type II membrane protein. Also exists a
CC        extracellular soluble form (By similarity).
CC    -!- PTM: The soluble form derives from the membrane form by
CC        proteolytic processing (By similarity).
CC    -!- SIMILARITY: Belongs to the tumor necrosis factor family.
CC    --------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a col
CC    between the Swiss Institute of Bioinformatics and the EMBL out
CC    the European Bioinformatics Institute. There are no restrictior
CC    use by non-profit institutions as long as its content is ir
CC    modified and this statement is not removed. Usage by and for c
CC    entities requires a license agreement (see http://www.isb-sib.ch/
CC    or send an email to license@isb-sib.ch).
CC    --------------------------------------------------------------
DR    EMBL; AF086711; AAD04375.1; -.
DR    HSSP; P29965; 1ALY.
DR    GO; GO:0016021; C:integral to membrane; ISS.
DR    GO; GO:0005174; F:CD40 receptor binding; ISS.
DR    GO; GO:0042100; P:B-cell proliferation; ISS.
DR    GO; GO:0006954; P:inflammatory response; ISS.
DR    GO; GO:0007159; P:leukocyte cell adhesion; ISS.
DR    GO; GO:0030168; P:platelet activation; ISS.
DR    InterPro; IPR003263; TNF_5.
DR    InterPro; IPR006052; TNF_family.
DR    InterPro; IPR008983; TNF_like.
DR    InterPro; IPR003636; TNF_subf.
DR    Pfam; PF00229; TNF; 1.
DR    PRINTS; PR01702; CD40LIGAND.
DR    ProDom; PD008600; TNF_5; 1.
DR    ProDom; PD002012; TNF_subf; 1.
DR    SMART; SM00207; TNF; 1.
DR    PROSITE; PS00251; TNF_1; 1.
DR    PROSITE; PS50049; TNF_2; 1.
KW    Cytokine; Transmembrane; Glycoprotein; Signal-anchor.
FT    CHAIN         1    260       TUMOR NECROSIS FACTOR LIGAND SUPERF
FT                                 MEMBER 5, MEMBRANE FORM.
FT    CHAIN       112    260       TUMOR NECROSIS FACTOR LIGAND SUPERF
FT                                 MEMBER 5, SOLUBLE FORM (BY SIMILARI
FT    DOMAIN        1     22       CYTOPLASMIC (POTENTIAL).
```

```
Db    165   AHLTINAASIPSGSH--KVTLSSWYHDRGWAKISMNT-----LSNGKLRVNQDGF

Qy    141   ---CQVHFDEGKAV--YLKLDLLV---NGVLALRCLEEFSATAASSPGPQLRL-
                 |||:|      |:|::    |:|: | ||  :| : :|:  |:|

Db    217   ANICFRHHETSGSVPTDYLQLMVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFY.
                 :|| :: |:|  :  |   | :  |: :|:  :|| || :| |:

Qy    189   SGLLPLRPGSSLRIRTLPWAHLKAAPFLYFGLFQV 224
                 :|| ||  |:|   ||:|      |||| ||  |:|

Db    277   GGFFKLRAGEISIQVSNPSLLDPDQDATYFGAFKV 312
```

FT   23   46   SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT   (POTENTIAL).
FT   47   260   EXTRACELLULAR (POTENTIAL).
FT   11   112   CLEAVAGE (BY SIMILARITY).
FT   77   217   POTENTIAL.
FT   39   239   N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   0 AA; 28688 MW; 604F69A19E98EB70 CRC64;

     7.8%; Score 90.5; DB 1; Length 260;
    28.1%; Pred. No. 0.86;
arity   16; Mismatches 53; Indels 31; Gaps 5;
onservative

Qy NSSSPLRYDRQIGEFT----------VIRAGLYYLYCQVHFDEGKAVYLKL 155
    |:.|:.| | :|.|| ||:|:|:||
    NPASVLRWAPK-GYTISSNLVSLENGKOLAVKRQGLYYVYAQVTFCSNRAASSQA 186

?   NGVLALRCLEEFSAT----AASSPGPQLRLCQVS----GLLPLRPGSSLRIRTL 205
    | || ||||| |||:|:||
    ---VASLCLHSPSGTERVLLRAASSRGSSKPCGQQSIHLGGVFELHPGASVFVNVT 241

?   ILKAAPFLTYFGLFQV 224
    : ||| ::
)VSHGTGFTSFGLLKL 260

STANDARD;   PRT;   318 AA.

9;
(Rel. 41, Created)
(Rel. 41, Last sequence update)
(Rel. 41, Last annotation update)
s factor ligand superfamily member 11 (Receptor activator
actor kappa B ligand) (RANKL) (TNF-related activation-
cine) (TRANCE) (Osteoprotegerin ligand) (OPGL) (Osteoclast
on factor) (ODF)
ANKL OR TRANCE OR OPGL.
jicus (Rat).
atazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
theria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
)116;

4 N.A.
l bone;
)945; PubMed=11092398;
J.K., Huang L., Gao X.H., Laird R., Liu D., Wysocki S.,
quence and functional characterization of the rat
receptor activator of NF-kB ligand.";
c. Res. 15:2178-2186(2000).

266-318 FROM N.A.
er 344;
2371; PubMed=11804028;
Kim N., van Wesenbeeck L., MacKay C., Mason-Savas A.,
Popoff S.N., Lengner C., van-Hul W., Choi Y.,
r.;
at the rat osteopetrotic mutation toothless (tl) is not in
(TRANCE, RANKL, ODF, OPGL) gene.";
Biol. 45:853-859(2001).
A/RANK. Osteoclast differentiation and activation factor.
r.; Cytokine that binds to TNFRSF11B/OPG and to
A/RANK. Osteoclast differentiation and activation factor.
the ability of dendritic cells to stimulate naive T-cell
ation. May be an important regulator of interactions
T cells and dendritic cells and may play a role in the
on of the T cell-dependent immune response. May also play
tant role in enhanced bone-resorption in humoral
cemia of malignancy.
Homotrimer (By similarity).
LAR LOCATION: Type II membrane protein and secreted (By
ty).
PECIFICITY: Highly expressed in thymus and bone tissues.
soluble form derives from the membrane form by

CC   proteolytic processing (By similarity).
CC   -!- SIMILARITY: Belongs to the tumor necrosis factor family.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a coll
CC   between the Swiss Institute of Bioinformatics and the EMBL out
CC   the European Bioinformatics Institute. There are no restriction
CC   use by non-profit institutions as long as its content is ir
CC   modified and this statement is not removed. Usage by and for c
CC   entities requires a license agreement (See http://www.isb-sib.ch/
CC   or send an email to license@isb-sib.ch).

DR   EMBL; AF187319; AAG17031.1; -.
DR   EMBL; AF425669; AAL23963.1; -.
DR   HSSP; P50591; 1DG.
DR   InterPro; IPR006052; TNF_family.
DR   InterPro; IPR009983; TNF_like.
DR   InterPro; IPR003636; TNF_subf.
DR   Pfam; PF00229; TNF; 1.
DR   ProDom; PD002012; TNF_subf; 1.
DR   SMART; SM00207; TNF; 1.
DR   PROSITE; PS00251; TNF 1; FALSE_NEG.
DR   PROSITE; PS50049; TNF 2; 1.
KW   Cytokine; Differentiation; Receptor; Glycoprotein; Transmembrane
KW   Signal-anchor.
FT   CHAIN   1   318   TUMOR NECROSIS FACTOR LIGAND SUPERF/
FT   MEMBER 11, MEMBRANE FORM.
FT   CHAIN   141   318   TUMOR NECROSIS FACTOR LIGAND SUPERF/
FT   MEMBER 11, SOLUBLE FORM.
FT   DOMAIN   1   47   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   48   68   SIGNAL-ANCHOR (TYPE-II MEMBRANE PRO'
FT   (POTENTIAL).
FT   DOMAIN   69   318   EXTRACELLULAR (POTENTIAL).
FT   SITE   140   141   CLEAVAGE (BY SIMILARITY).
FT   CARBOHYD 199   199   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD 264   264   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT 317   317   I -> M (IN REF. 2).
SQ   SEQUENCE 318 AA; 35370 MW; 4B87A4D706AD098F CRC64;

Query Match   7.8%; Score 90.5; DB 1; Length 318;
Best Local Similarity 23.3%; Pred. No. 1.1;
Matches   64; Conservative 36; Mismatches 108; Indels 67;

Qy   6 LALACLGLLLVVVSLGSWATLSAQ-EPSQEELTABDR----------REPPELNP
         :.| ||::| |: .| .|| .||:
Db   51 LALLGLGLGQVVCSTALFLYFRAQMDPNR---ISEDSTRCFYRIRLRENTGLQD

Qy   55 QD----------VVPFLEQLVRPRRSA----------PKGRKARPR
         :| :| || ::
Db 108 EDTEALPDSCRRMKQAFQGAVQRELQHIVGPQRFSGVPAMMEGSWLDVARRGKPE

Qy   86 HYEVHPRPGQDGAQAGVDGTVSGWEE---TKINSSSPLRYDRQIGEFTVIRAGL
         ||:   :: :: :| |||:|:||
Db 168 HLTINAADIPSGSH---KVSLSSWYHDRGWAKISNMT----LSNGKLRVNQDGF

Qy 142 QV----HFDEGK-AVYLKIDLLV---NGVLALRCLEEFSATRAASSPGPQLRL-
         |: |.|:| .:|:| .|| |:| |:
Db 220 NICFRHHETSGSVPADYLQLMVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFY

Qy 190 GLLPLRPGSSLRIRTLPWAHLKAAPFLTYFGLFQV 224
         | |: | : . || ||:: ||||:||:|
Db 280 GFFKIRAGEEISVQVSNPSLLDPDQDATYFGAFKV 314

RESULT 7
PPBN_HUMAN
ID   PPBN_HUMAN   STANDARD;   PRT;   532 AA.
AC   P10696; Q16727; Q9CCM1;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Alkaline phosphatase, placental-like precursor (EC 3.1.3.1) (Nag
DE   isozyme) (Germ-cell alkaline phosphatase) (GCAP) (PLAP-like) (AL
GN   ALPPL2 OR ALPPL.

CC   (liver/bone/kidney).
CC   -!- SIMILARITY: Belongs to the alkaline phosphatase family.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a colla
CC   between the Swiss Institute of Bioinformatics and the EMBL outs
CC   the European Bioinformatics Institute. There are no restriction
CC   use by non-profit institutions as long as its content is in
CC   modified and this statement is not removed. Usage by and for cc
CC   entities requires a license agreement (See http://www.isb-sib.ch/a
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; X55958; CAA39425.1; -.
DR   EMBL; J04948; AAA51700.1; -.
DR   EMBL; J03252; AAA98616.1; -.
DR   EMBL; X53279; CAA37374.1; -.
DR   EMBL; BC014139; AAH14139.1; -.
DR   EMBL; X07247; CAA30232.1; ALT_SEQ.
DR   PIR; S12076; S12076.
DR   HSSP; P00634; 1AJC.
DR   Siena-2DPAGE; P10696; -.
DR   Genew; HGNC-441; ALPPL2.
DR   MIM; 171810; -.
DR   GO; GO:0016020; C:membrane; NAS.
DR   GO; GO:0004035; F:alkaline phosphatase activity; NAS.
DR   GO; GO:0016310; P:phosphorylation; NAS.
DR   InterPro; IPR001952; Alk_phosphtse.
DR   Pfam; PF00245; alk_phosphatase; 1.
DR   PRINTS; PR00113; ALKPHPHTASE.
DR   ProDom; PD001868; Alk_phosphtse; 1.
DR   SMART; SM00098; alkPPc; 1.
DR   PROSITE; PS00123; ALKALINE_PHOSPHATASE; 1.
KW   Hydrolase; zinc; Magnesium; Phosphorylation; Transmembrane;
KW   Multigene family; Glycoprotein; GPI-anchor; Signal.
FT   SIGNAL   1   19   POTENTIAL.
FT   CHAIN   20   503   ALKALINE PHOSPHATASE, PLACENTAL-LIKE
FT   PROPEP   504   532   REMOVED IN MATURE FORM.
FT   DISULFID   140   202   BY SIMILARITY.
FT   DISULFID   486   493   BY SIMILARITY.
FT   ACT_SITE   111   111   PHOSPHOSERINE INTERMEDIATE.
FT   CARBOHYD   141   141   N-LINKED (GLCNAC...)   (POTENTIAL).
FT   CARBOHYD   268   268   N-LINKED (GLCNAC...)   (POTENTIAL).
FT   CONFLICT   57   57   I -> M (IN REF. 3).
FT   CONFLICT   152   152   V -> M (IN REF. 2, 4, 5 AND 6).
FT   CONFLICT   178   178   H -> A (IN REF. 4 AND 5).
FT   CONFLICT   260   260   H -> R (IN REF. 5).
FT   CONFLICT   273   273   L -> M (IN REF. 4 AND 5).
FT   CONFLICT   316   316   L -> R (IN REF. 1, 4 AND 5).
FT   CONFLICT   380   380   V -> L (IN REF. 2).
FT   CONFLICT   498   498   P -> R (IN REF. 2 AND 5).
FT   CONFLICT   498   498   P -> S (IN REF. 4).
FT   CONFLICT   531   531   A -> T (IN REF. 4).
SQ   SEQUENCE   532 AA; 57315 MW; 84AB5B28F13D6DB2 CRC64;

Query Match   7.6%; Score 88; DB 1; Length 532;
Best Local Similarity 28.6%; Pred. No. 3.4;
Matches   44; Conservative 16; Mismatches 44; Indels 50; G

Qy   76 KARPRRAIAAHYEV--HPRPG--QDGAQAGVDGTVSGWEETKINSSSSPLRYDRQ
           ||||:|| ::||| ::||::|||:||||:||||||
Db   400 KARDRKA----YTVLLYGNGPGYVLKDGARPDVTESESGSPEYRQQSAVPLDGETH

Qy   130 TVIRAGLYYLYCQVHFDEGKAVYLKDLLVNGV------LALRCLEEFSA--
           ||:|| | || ||| : ||| |:|
Db   456 VAV-----FARGPQAH----LVHGVQEQTFIAHVMAFAACLEPYTACD

Qy   174 ---TAASSPQLRLCQVSGLLPLRPGSSLRIRT 204
              :|| | | | : ||||||| :| | | | :
Db   499 AGTTDAAHPGPSV---VPAILPLLAGTLLLLGT 528

RESULT 8
MINT_HUMAN
ID   MINT_HUMAN   STANDARD;   PRT;   3664 AA.

Human).
...zoa; Chordata; Craniata; Vertebrata; Euteleostomi;
...eria; Primates; Catarrhini; Hominidae; Homo.
...5;

N.A.
...79; PubMed=2162249;
...uss A.W.;
...a Nagao-type, phosphatidylinositol-glycan anchored
...hatase in human choriocarcinomas.";
...:3956-3962(1990).

N.A.
...arcinoma;
...96; PubMed=2745460;
...atanabe T., Li W.I., Soong B.-W., Chou J.Y.;
the germ cell alkaline phosphatase gene in human
...a cells.";
264:12611-12619(1989).

N.A.
...32; PubMed=2834730;
...anes T.;
ved Nagao isozyme is encoded by a germ-cell alkaline
ne.";
ad. Sci. U.S.A. 85:3024-3028(1988).

N.A.

11; PubMed=2297757;
s J.W., Sack T.L., Kim Y.S.;
ning of complementary DNAs encoding alkaline
human colon cancer cells.";
:1085-1091(1990).

N.A.

a;
57; PubMed=12477932;
..., Feingold E.A., Grouse L.H., Derge J.G.,
Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
.. Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
uellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
Ewan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
orley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
on E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
dan A., Young A.C., Shevchenko Y., Bouffard G.G.,
, Touchman J.W., Green E.D., Dickson M.C.,
, Grimwood J., Schmutz J., Myers R.M.,
S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
schein J.E., Jones S.J.M., Marra M.A.;
nd initial analysis of more than 15,000 full-length
se cDNA sequences.";
ad. Sci. U.S.A. 99:16899-16903(2002).

-157 FROM N.A.

78; PubMed=3387245;
1 H., Kan Y.W., Kam W.;
a sequence of a putative human placental alkaline
ke gene.";
Res. 16:5694-5694(1988).
ACTIVITY: An orthophosphoric monoester + H(2)O = an
phosphate.
...homodimer.
AR LOCATION: Attached to the membrane by a GPI-anchor.
ECIFICITY: TRACE AMOUNTS IN THE TESTIS AND THYMUS,
EVATED AMOUNTS IN GERM CELL TUMORS.
EOUS: In most mammals there are four different isozymes:
, placental-like, intestinal and tissue non-specific

```
AC   ...8; Q9NWH5; Q9UQ01; Q9Y556;
DT   ...(Rel. 42, Created)
DT   ...(Rel. 42, Last sequence update)
DT   ...(Rel. 42, Last annotation update)
DE   ...ing protein (SMART/HDAC1 associated repressor protein).
OS   ... OR KIAA0929.
OS   ...(Human)
OC   ...tazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   ...theria; Primates; Catarrhini; Hominidae; Homo.
OX   ...506;
RN
RP   ...N.A., FUNCTION, SUBCELLULAR LOCATION, INDUCTION,
RP   ...AND INTERACTION WITH NCOR2; HDAC1; HDAC2; RBBP4; MBD3;
RP   ...1.
RC   ..., and Pituitary;
RX   ...190; PubMed=11331609;
RA   ...es M., Xie W., Kao H.-Y., Ordentlich P., Tsai C.-C.,
RA   ...s R.M.;
RT   ...ducible cofactor that integrates nuclear receptor
RT   ...d activation.";
RL   ...5:1140-1151(2001).
RN
RP   ...N.A.
RX
RL   ...JN-2003) to the EMBL/GenBank/DDBJ databases.
RN
RP   ...294-3664 FROM N.A.
RA   ...ickle E.;
RL   ...JL-1999) to the EMBL/GenBank/DDBJ databases.
RN
RP   ...793-1595 FROM N.A., AND VARIANT PRO-1091.
RC   ...o, and Teratocarcinoma;
RA   ...ca T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   ..., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA   ...omura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA   ...pekura T., Ishii S., Kawai Y., Saito Y., Yamamoto J.,
RA   ..., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;
RT   ...cDNA sequencing project.";
RL   ...JG-2000) to the EMBL/GenBank/DDBJ databases.
RN
RP   ...2002-3664 FROM N.A.
RX
RA   ...5063; PubMed=10231032;
RA   ...shikawa K.-I., Suyama M., Kikuno R., Hirosawa M.,
RA   ...Tanaka A., Kotani H., Nomura N., Ohara O.;
RT   ...of the coding sequences of unidentified human genes. XIII.
RT   ...sequences of 100 new cDNA clones from brain which code
RT   ...oteins in vitro.";
RL   ...3-70(1999).
RN
RP   ...WITH PPARD.
RX   ...4127; PubMed=11867749;
RA   ...M., Evans R.M.;
RT   ...ome proliferator-activated receptor delta, an integrator
RT   ...tional repression and nuclear receptor signaling.";
RL   ...Acad. Sci. U.S.A. 99:2613-2618(2002).
RN
RP   ...D INTERACTION WITH RBPSUH.
RX   ...1914; PubMed=12374742;
RA   ...ostezka U., Astrahantseff K., Bourteele S., Dillinger K.,
RA   ...Ludwig L., Wilda M., Hameister H., Knoechel W., Liptay S.,
RT   ...novel component of the Notch/RBP-Jkappa signalling
RL   ...417-5426(2002).
RN
RP   ...LLOGRAPHY (1.8 ANGSTROMS) OF SPOC DOMAIN.
RX   ...7836; PubMed=12897056;
RA   ...Schwabe J.W.R.;
RT   ...structural motif reveals the essential transcriptional
RT   ...unction of Spen proteins and their role in developmental
RL   ...7:1909-1920(2003).
```

```
CC   -!- FUNCTION: Essential corepressor protein, which probably regu
CC       different key pathways such as the Notch pathway. Negative
CC       regulator of the Notch pathway via its interaction with RBPS
CC       which prevents the association between NOTCH1 and RBPSUH, an
CC       therefore suppresses the transactivation activity of Notch
CC       signaling. Blocks the differentiation of precursor B cells i
CC       marginal zone B cells. Probably represses transcription via t
CC       recruitment of large complexes containing histone deacetylas
CC       proteins. May bind both to DNA and RNA.
CC   -!- SUBUNIT: Interacts with MSX2 (By similarity). Interacts with
CC       NCOR2, HDAC1, HDAC2, RBBP4, MBD3 and MTA1L1. Interacts with
CC       RBPSUH; this interaction may prevent the interaction between
CC       RBPSUH and NOTCH1. Interacts with the nuclear receptors RAR
CC       PPARD. Interacts with RAR in absence of ligand. Bind to the
CC       steroid receptor RNA coactivator SRA.
CC   -!- SUBCELLULAR LOCATION: Nuclear. Associates with chromatin.
CC   -!- TISSUE SPECIFICITY: Expressed at high level in brain, testis,
CC       spleen and thymus. Expressed at intermediate level in kidney,
CC       liver, mammary gland and skin.
CC   -!- INDUCTION: By hormone 17-beta-estradiol (E2).
CC   -!- DOMAIN: The RID domain mediates the interaction with nuclear
CC       receptors (By similarity).
CC   -!- DOMAIN: The SPOC domain, which mediates the interaction with
CC       NCOR2, is essential for the repressive activity.
CC   -!- SIMILARITY: Belongs to the Spen family.
CC   -!- SIMILARITY: Contains 1 RID (receptor interacting) domain.
CC   -!- SIMILARITY: Contains 4 RNA recognition motif (RRM) domains.
CC   -!- SIMILARITY: Contains 1 SPOC domain.
CC   -!- CAUTION: Ref.2 sequences differ from that shown due to errou
CC       gene model prediction.
CC   -----------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a col
CC   between the Swiss Institute of Bioinformatics and the EMBL ou
CC   the European Bioinformatics Institute. There are no restrictio
CC   use by non-profit institutions as long as its content is i
CC   modified and this statement is not removed. Usage by and for
CC   entities requires a license agreement (See http://www.isb-sib.ch).
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------
DR   EMBL; AF356524; AAK52750.1; -.
DR   EMBL; AL034555; CAB85442.1; ALT_SEQ.
DR   EMBL; AL034555; CAB85444.1; ALT_SEQ.
DR   EMBL; AL450998; -; NOT_ANNOTATED_CDS.
DR   EMBL; AL096858; CAB51072.1; ALT_INIT.
DR   EMBL; AK000882; BAA91405.1; ALT_INIT.
DR   EMBL; AK022949; BAB14324.1; ALT_INIT.
DR   EMBL; AB023146; BAA76773.1; -.
DR   InterPro; IPR000504; RNA_rec_mot.
DR   PDB; 1OW1; 19-AUG-03.
DR   Pfam; PF00076; rrm; 4.
DR   SMART; SM00360; RRM; 4.
DR   PROSITE; PS50102; RRM; 4.
DR   PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
DR   PROSITE; PS50917; SPOC; 1.
KW   Transcription regulation; Repressor; Nuclear protein; DNA-bindin
KW   Repeat; RNA-binding; Coiled coil; 3D-structure; Polymorphism.
FT   DOMAIN       1    573       DNA-BINDING (BY SIMILARITY).
FT   DOMAIN       6     81       RNA-BINDING (RRM) 1.
FT   DOMAIN     335    415       RNA-BINDING (RRM) 2.
FT   DOMAIN     438    513       RNA-BINDING (RRM) 3.
FT   DOMAIN     517    589       RNA-BINDING (RRM) 4.
FT   DOMAIN     688    715       COILED COIL (POTENTIAL).
FT   DOMAIN     977   1004       COILED COIL (POTENTIAL).
FT   DOMAIN    1170   1191       COILED COIL (POTENTIAL).
FT   DOMAIN    1408   1428       COILED COIL (POTENTIAL).
FT   DOMAIN    1496   1529       COILED COIL (POTENTIAL).
FT   DOMAIN    1592   1612       COILED COIL (POTENTIAL).
FT   DOMAIN    1928   1944       COILED COIL (POTENTIAL).
FT   DOMAIN    2201   2707       RID.
FT   DOMAIN    3498   3664       SPOC.
FT   DOMAIN    2130   2464       INTERACTION WITH MSX2 (BY SIMILARIT
FT   DOMAIN    2709   2870       INTERACTION WITH RBPSUH (BY SIMILAR
FT   DOMAIN     125    277       ARG-RICH.
```

```
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RA   Rowen L., Madan A., Qin S., Shaffer T., James R., Ratcliffe A.,
RA   Abbasi N., Dickhoff R., Loretz C., Madan A., Dors M., Young J.,
RA   Lasky S., Hood L.;
RT   "Sequence of the human major histocompatibility complex class III
RT   region.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RA   Shiina S., Tamiya G., Oka A., Inoko H.;
RT   "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I regior
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANTS GLU-70 AND
RP   PRO-111.
RA   Rieder M.J., Armel T.Z., Carrington D.P., Chung M.-W., Lee K.L.,
RA   Poel C.L., Toth E.J., Yi Q., Nickerson D.A.;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANTS ARG-84 AND PHE-87.
RA   Rieder M.J., Livingston R.J., Daniels M.R., Montoya M.A., Chung M
RA   Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D
RA   Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: Cytokine that binds to LTBR/TNFRSF3. May play a spe
CC       role in immune response regulation. Provides the membrane anc
CC       for the attachment of the heterotrimeric complex to the cell
CC       surface. Isoform 2 is probably non-functional.
-!- SUBUNIT: Heterotrimer of either two LTB and one LTA subunits
CC       (less prevalent) one LTB and two LTA subunits.
-!- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
-!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=Q06643-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q06643-2; Sequence=VSP_006441, VSP_006442;
-!- TISSUE SPECIFICITY: Spleen and thymus.
-!- SIMILARITY: Belongs to the tumor necrosis factor family.
```

```
DR   EMBL; L11016; AAA99888.1; -.
DR   EMBL; U89922; AAC51769.1; -.
DR   EMBL; U79029; AAB37342.1; -.
DR   EMBL; L11015; AAA36191.1; -.
DR   EMBL; Y14768; CAA75069.1; -.
DR   EMBL; AF129756; AAD18089.1; -.
DR   EMBL; AP000505; BAB63395.1; -.
DR   EMBL; AY070219; AAL49954.1; -.
DR   EMBL; AY070219; AAL49955.1; -.
DR   EMBL; AY216497; AAO21134.1; -.
DR   PIR; A46066; A46066.
DR   PIR; JC5645; JC5645.
DR   HSSP; P01374; 1TNR.
DR   Genew; HGNC:6711; LTB.
DR   MIM; 600978; -.
DR   GO; GO:0005102; F:receptor binding; TAS.
DR   GO; GO:0015070; F:toxin activity; NAS.
DR   GO; GO:0007267; P:cell-cell signaling; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR006053; TNF abc.
DR   InterPro; IPR006052; TNF family.
DR   InterPro; IPR008983; TNF_like.
DR   InterPro; IPR003636; TNF_subf.
DR   Pfam; PF00229; TNF; 1.
```

```
0    325    SER-RICH.
6    810    ARG-RICH.
4    697    TYR-RICH.
8    2520   PRO-RICH.
0    3482   PRO-RICH.
0    970    A -> V (in dbSNP:848208).
             /FTId=VAR_017119.
1    1091   L -> P (in dbSNP:848209).
             /FTId=VAR_017120.
0    2360   N -> D (in dbSNP:848210).
             /FTId=VAR_017121.
6    956    G -> D (IN REF. 4).
4 AA;   402245 MW;   5228C58533E5B27B CRC64;

rity       7.6%;   Score 88;   DB 1;   Length 3664;
         23.1%;   Pred. No. 35;
nservative 20; Mismatches  62; Indels  58; Gaps  7;

'AEDRREPPELNPQTEESQ----------------DVVPFLEQLVRPRRSAPK----- 73
 ||:|:||:|:|        :::||||||
'AGPANRPPEPHTQVQRAQAETGPTSFPSPVSVSMKPDL-PVSLPTQTAPKQPLFV 3459

'KARPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSSSPLRYDRQIGEFT 130
 |||||||        |||    :||||
'PSTPPGLVLPHTEFQPAPKQDSS--------------PHLTSQRPVDMVQLLKKYP 3506

'LYYL-----YCQVHFDEGKAVYLKLDLLVNGVLALRCLEEFSATAASSPGQLRL 185
 :::|          |||||
'LLALKNDTAAVQLHFVSG----------NNVLAHRSL------PLSEGGPPLRI 3549

37

651
```

```
STANDARD;      PRT;     244 AA.
); Q99761;
Rel. 29, Created)
Rel. 29, Last sequence update)
Rel. 43, Last annotation update)
ta (LT-beta) (Tumor necrosis factor C) (TNF-C) (Tumor
or ligand superfamily member 3).
OR TNFC.
(Human)
azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
neria; Primates; Catarrhini; Hominidae; Homo.
06;

N.A. (ISOFORMS 1 AND 2).

381; PubMed=7916655;
Ngam-Ek A., Lawton P., Demarinis J., Tizard R.,
sion C., O'Brine-Greco B., Foley S.F., Ware C.F.;
ta, a novel member of the TNF family that forms a
omplex with lymphotoxin on the cell surface.";
56(1993).

N.A. (ISOFORMS 1 AND 2).
965; PubMed=9294492;
Renard N., Charlot C., Bienvenu J., Coiffier B.,

on of two lymphotoxin beta isoforms expressed in human
lines and non-Hodgkin's lymphomas.";
hys. Res. Commun. 238:273-276(1997).

N.A. (ISOFORM 1).
Milner C.M., Campbell R.D.;
of the immunoglobulin superfamily and a V-ATPase G
mongst the predicted products of novel genes close to the
the human MHC.";
P-1997) to the EMBL/GenBank/DDBJ databases.
```

RT   duplication blocks, diversity and complexity from shark to man."
RL   Immunol. Rev. 190:95-122(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22709134; PubMed=12799463;
RA   Anzai T., Shiina T., Kimura N., Yanagiya K., Kohara S., Shigenar
RA   Yamagata T., Kulski J.K., Naruse T.K., Fujimori Y., Fukurumi Y.,
RA   Yamazaki M., Tashiro H., Iwamoto C., Umehara Y., Imanishi T.,
RA   Meyer A., Ikeo K., Gojobori T., Bahram S., Inoko H.;
RT   "Comparative sequencing of human and chimpanzee MHC class I regio
RT   unveils insertions/deletions as the major path to genomic
RT   divergence.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:7708-7713(2003).
CC   -!- FUNCTION: Cytokine that binds to LTBR/TNFRSF3. May play a spe
CC   role in immune response regulation. Provides the membrane an
CC   for the attachment of the heterotrimeric complex to the cell
CC   surface (By similarity).
CC   -!- SUBUNIT: Heterotrimer of either two LTB and one LTA subunits
CC   (less prevalent) two LTA and one LTB subunits (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
CC   -!- SIMILARITY: Belongs to the tumor necrosis factor family.
CC
DR   EMBL; AB054536; BAB83881.1; -.
DR   EMBL; AB100082; BAC78156.1; -.
DR   InterPro; IPR006053; TNF_abc.
DR   InterPro; IPR006052; TNF_family.
DR   InterPro; IPR008983; TNF_like.
DR   InterPro; IPR003636; TNF_subf.
DR   Pfam; PF00229; TNF; 1.
DR   PRINTS; PR01234; TNECROSISFCT.
DR   ProDom; PD002012; TNF_subf; 1.
DR   SMART; SM00207; TNF; 1.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 1.
KW   Cytokine; Transmembrane; Glycoprotein; Signal-anchor.
FT   DOMAIN   1   18   CYTOPLASMIC (POTENTIAL)
FT   TRANSMEM   19   48   SIGNAL-ANCHOR (TYPE II MEMBRANE PRO
FT   (POTENTIAL).
FT   DOMAIN   49   244   EXTRACELLULAR (POTENTIAL)
FT   CARBOHYD   222   222   N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   244 AA;   25420 MW; A404785833SD5B97 CRC64;

Query Match   7.5%; Score 87; DB 1; Length 244;
Best Local Similarity 23.0%; Pred. No. 1.6;
Matches   64; Conservative 26; Mismatches 80; Indels 108;

Qy   3 SLGLALA--CLGLLLVVSLGSWATLSAQEPSQEELTAEDR----------R
Db   18 SLLLAVAGATSLVTLLLAVPITVLAVLALVPQDQGGLVTETADPGAQQGLGFQ

Qy   48 NPQTEESQDVVPFLEQLVRPRRSAPKGRKARPRRAIAAHYEVHPRPGQDGAQGV
Db   78 EPET----DLSPGLP----------------AAHLIGAPLKGQG---

Qy   108 GWEETKINS--SSPLRYDRQIGEFTVIRAGLYYLYCQVHF-------DEGKA
Db   103 GWETTKEQAFLTSGTQFSDAEG-LALPQDGLYYLYCLVGYRGRTPPGGDPQGRS

Qy   156 DLLVNGVLALRCLEEFSATAASPG-PQLRL----------------
Db   162 SL-------------YRAGGAYGPGTPELLLEGAETVTPVLDPARRQYGPLWYT

Qy   190 GLLPLRPGSSLRIRTLPWAHLKAAPFL---TYFGLFQV 224
Db   209 GLVQLRRGERVYVNI---SHPDMVDFARGKTFFGAVMV 243

;34; TNECROSISFCT.
012; TNF_subf; 1.
7; TNF; 1.
251; TNF_1; 1.
049; TNF_2; 1.
nsmembrane; Glycoprotein; Signal-anchor;
plicing; Polymorphism.
            1   18   CYTOPLASMIC (POTENTIAL).
           19   48   SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
                        (POTENTIAL).
           49   244   EXTRACELLULAR (POTENTIAL).
          222   222   N-LINKED (GLCNAC...) (POTENTIAL).
           53   77   GLVTETADPGAQAQQGLGFQKLPEE -> GLGFRSCQRRSQ
                        KQISAPGSQLPTS (in isoform 2).
                        /FTId=VSP_006441.
           78   244   Missing (In isoform 2).
                        /FTId=VSP_006442.
           70   70   G -> E.
                        /FTId=VAR_013025.
           84   84   S -> R.
                        /FTId=VAR_016331.
           87   87   L -> F.
                        /FTId=VAR_016332.
          11   111   A -> P.
                        /FTId=VAR_013026.
           60   69   DPGAQAQQGL -> GLSAPGSGRT (IN REF. 2;
                        AAB37342).
4 AA;   25390 MW; F41564598300ED4C CRC64;

arity   7.5%; Score 87; DB 1; Length 244;
onservative 26; Mismatches 80; Indels 108; Gaps 13;

ALA--CLGLLLVVSLGSWATLSAQEPSQEELTAEDR----------REPPEL 47
AVAGATSLVTLLLAVPITVLAVLALVPQDQGGLVTETADPGAQQGLGFQKLPEE 77

EESQDVVPFLEQLVRPRRSAPKGRKARPRRAIAAHYEVHPRPGQDGAQGVDGTVS 107
---DLSPGLP--------------------AAHLIGAPLKGQG---------L 102

ETKINS--SSPLRYDRQIGEFTVIRAGLYYLYCQVHF---------DEGKAVYLKL 155
TTKEQAFLTSGTQFSDAEG-LALPQDGLYYLYCLVGYRGRAPPGGGDPQGRSVTLRS 161

JNGVLALRCLEEFSATAASPG-PQLRL----------------------CQVS 189
---------YRAGGAYGPGTPELLLEGAETVTPVLDPARRQGYGPLWYTSVGFG 208

PLRPGSSLRIRTLPWAHLKAAPFL---TYFGLFQV 224
QLRRGERVYVNI---SHPDMVDFARGKTFFGAVMV 243

STANDARD;   PRT;   244 AA.

(Rel. 43, Created)
(Rel. 43, Last sequence update)
(Rel. 43, Last annotation update)
beta (LT-beta) (Tumor necrosis factor C) (TNF-C) (Tumor
tor ligand superfamily member 3).
3 OR TNFC.
tes (Chimpanzee).
etazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
theria; Primates; Catarrhini; Hominidae; Pan.
598;
M.N.A.
1002; PubMed=12493009;
Shiina T., Anzai T., Kohara S., Inoko H.;
genomic analysis of the MHC: the evolution of class I

```
FT   DISULFID     178    218       POTENTIAL.
FT   CARBOHYD     240    240       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE     261 AA;   29357 MW;   85E1588B507901B5 CRC64;

Query Match                    7.5%;  Score 87;  DB 1;  Length 261;
Best Local Similarity         25.9%;  Pred. No. 1.8;
Matches   35;  Conservative   21;  Mismatches   55;  Indels   24;  Ga

Qy  105 TVSGWEE---TKINSSSPLRYDRQIGFTVIRAGLYYLYCQVHFDEGKAVYLKLDI
            :   :::  ::   : : |||:||:::  ::|| :|||:|: ::
Db  136 SVLQWAERGYYTMSNNLVTLENGKQL---TVKRQGLYYIYAQVTFCSNREASSQAPF

Qy  161 GVLALRCLEEF-----SATAASSPGPQLRLC----QVSGLLPLRPGSSLRIRTLF
            :|:|:  |  ||||:  :::|:|:::||:|
Db  193 --LCLKPNRFERILLRAANTHSSAKP---CGQQSIHLGGIFELQPGASVFVNVTI

Qy  210 LKAAPFLIYFGLFQV 224
            : ::   ||| |
Db  247 VSHGTGFTSFGLLKL 261
```

```
RESULT 12
TNF5_CALJA
ID   TNF5_CALJA     STANDARD;      PRT;   261 AA.
AC   Q9BDN3;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Tumor necrosis factor ligand superfamily member 5 (CD40 ligand) (C
DE   L) (CD154 protein).
GN   TNFSF5 OR CD40LG OR CD40L.
OS   Callithrix jacchus (Common marmoset).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
OC   Callithrix.
OX   NCBI_TaxID=9483;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=lymphocytes;
RX   MEDLINE=21383618; PubMed=11491535;
RA   Villinger F., Bostik P., Mayne A.E., King C.L., Genain C.P.,
RA   Weiss W.R., Ansari A.A.;
RT   "Cloning, sequencing, and homology analysis of nonhuman primate
RT   Fas/Fas-ligand and co-stimulatory molecules.";
RL   Immunogenetics 53:315-328(2001).
CC   -!- FUNCTION: Cytokine that binds to TNFRSF5. Mediates B-cell
CC       proliferation in the absence of co-stimulus as well as IgE
CC       production in the presence of IL-4. Involved in immunoglobulin
CC       class switching (By similarity)
CC   -!- SUBUNIT: Homotrimer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein. Also exists as
CC       extracellular soluble form (By similarity).
CC   -!- PTM: The soluble form derives from the membrane form by
CC       proteolytic processing (By similarity).
CC   -!- SIMILARITY: Belongs to the tumor necrosis factor family.
CC   ----
CC   This SWISS-PROT entry is copyright. It is produced through a colla
CC   between the Swiss Institute of Bioinformatics and the EMBL outs
CC   the European Bioinformatics Institute. There are no restriction
CC   use by non-profit institutions as long as its content is in
CC   modified and this statement is not removed. Usage by and for co
CC   entities requires a license agreement (See http://www.isb-sib.ch/
CC   or send an email to license@isb-sib.ch).
CC   ----
DR   EMBL; AF344844; AAK37603.1; -.
DR   HSSP; P29965; 1ALY.
DR   GO; GO:0016021; C:integral to membrane; ISS.
DR   GO; GO:0005174; F:CD40 receptor binding; ISS.
DR   GO; GO:0042100; P:B-cell proliferation; ISS.
DR   GO; GO:0006954; P:inflammatory response; ISS.
DR   GO; GO:0007159; P:leukocyte cell adhesion; ISS.
DR   GO; GO:0030168; P:platelet activation; ISS.
DR   InterPro; IPR003263; TNF_5.
```

```
STANDARD;      PRT;   261 AA.
el. 41, Created)
el. 41, Last sequence update)
el. 41, Last annotation update)
factor ligand superfamily member 5 (CD40 ligand) (CD40-
ein).
LG OR CD40L.
tus (Night monkey) (Dourocouli).
azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
eria; Primates; Platyrrhini; Cebidae; Aotinae; Aotus.
5;

N.A.
ytes;
18; PubMed=11491535;
Bostik P., Mayne A.E., King C.L., Genain C.P.,
sari A.A.;
encing, and homology analysis of nonhuman primate
and co-stimulatory molecules.";
53:315-328(2001).
Cytokine that binds to TNFRSF5. Mediates B-cell
ion in the absence of co-stimulus as well as IgE
in the presence of IL-4. Involved in immunoglobulin
ching (By similarity).
omotrimer (By similarity).
R LOCATION: Type II membrane protein. Also exists as an
lar soluble form (By similarity).
oluble form derives from the membrane form by
c processing (By similarity).
: Belongs to the tumor necrosis factor family.
----
T entry is copyright. It is produced through a collaboration
wiss Institute of Bioinformatics and the EMBL outstation -
ioinformatics Institute. There are no restrictions on its
rofit institutions as long as its content is in no way
his statement is not removed. Usage by and for commercial
res a license agreement (See http://www.isb-sib.ch/announce/
il to license@isb-sib.ch).
----
; AAK37542.1; -.
1ALY.
; C:integral to membrane; ISS.
; F:CD40 receptor binding; ISS.
; P:B-cell proliferation; ISS.
; P:inflammatory response; ISS.
; P:leukocyte cell adhesion; ISS.
; P:platelet activation; ISS.
03263; TNF_5.
06052; TNF_family.
08983; TNF_like.
03636; TNF_subf.
TNF; 1.
2; CD40LIGAND.
00; TNF_5; 1.
12; TNF_subf; 1.
; TNF; 1.
51; TNF_1; 1.
49; TNF_2; 1.
msmembrane; Glycoprotein; Signal-anchor.
1    261   TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
            MEMBER 5, MEMBRANE FORM.
3    261   TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
            MEMBER 5, SOLUBLE FORM (BY SIMILARITY).
1    22    CYTOPLASMIC (POTENTIAL).
3    43    SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
            (POTENTIAL).
4    261   EXTRACELLULAR (POTENTIAL).
2    113   CLEAVAGE (BY SIMILARITY).
```

```
RT   HVEM/TR2, stimulates proliferation of T cells and inhibits HT29 c
RT   growth.";
RL   J. Biol. Chem. 273:27548-27556(1998).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 2), AND PROCESSING.
RX   MEDLINE=21528948; PubMed=11673523;
RA   Granger S.W., Butrovich K.D., Houshmand P., Edwards W.R., Ware C.
RT   "Genomic characterization of LIGHT reveals linkage to an immune
RT   response locus on chromosome 19p13.3 and distinct isoforms genera
RT   by alternate splicing or proteolysis.";
RL   J. Immunol. 167:5122-5128(2001).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: Cytokine that binds to TNFRSF3/LTBR. Binding to the
CC       decoy receptor TNFRSF6B modulates its effects. Activates NFKB
CC       stimulates the proliferation of T cells, and inhibits growth
CC       the adenocarcinoma HT-29. Acts as a receptor for Herpes simpl
CC       virus.
CC   -!- SUBUNIT: Homotrimer.
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein and secreted
CC       (isoform 1); Cytoplasmic (isoform 2).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=O43557-1; Sequence=Displayed;
CC       Name=2; Synonyms=LIGHT delta-TM;
CC         IsoId=O43557-2; Sequence=VSP_006452;
CC   -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE SPLEEN BUT
CC       FOUND IN THE BRAIN. WEAKLY EXPRESSED IN PERIPHERAL LYMPHOID
CC       TISSUES AND IN HEART, PLACENTA, LIVER, LUNG, APPENDIX, AND KI
CC       AND NO EXPRESSION SEEN IN FETAL TISSUES, ENDOCRINE GLANDS, OF
CC       NONHEMATOPOIETIC TUMOR LINES.
CC   -!- INDUCTION: UPREGULATED AFTER T-CELL ACTIVATION.
CC   -!- PTM: N-glycosylated.
CC   -!- PTM: The soluble form of isoform 1 derives from the membrane
CC       by proteolytic processing.
CC   -!- SIMILARITY: Belongs to the tumor necrosis factor family.
CC   -!- CAUTION: Ref.4 sequence differs from that shown due to a
CC       frameshift in position 178.
CC   --------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a col
CC   between the Swiss Institute of Bioinformatics and the EMBL out
CC   the European Bioinformatics Institute. There are no restrictio
CC   use by non-profit institutions as long as its content is i
CC   modified and this statement is not removed. Usage by and for (
CC   entities requires a license agreement (See http://www.isb-sib.ch).
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------
DR   EMBL; AF036581; AAC39563.1; -.
DR   EMBL; AF064090; AAC25169.1; -.
DR   EMBL; AY028261; AAK26160.1; -.
```

```
006052; TNF_family.
008983; TNF_like.
003636; TNF_subf.
; TNF; 1.
02; CD40LIGAND.
600; TNF 5; 1.
012; TNF_subf; 1.
7; TNF; 1.
251; TNF 1; 1.
049; TNF_2; 1.
nsmembrane; Glycoprotein; Signal-anchor.
1    261   TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
           MEMBER 5. MEMBRANE FORM.
13   261   TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
           MEMBER 5, SOLUBLE FORM (BY SIMILARITY).
1    22    CYTOPLASMIC (POTENTIAL).
23   43    SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
           (POTENTIAL).
44   261   EXTRACELLULAR (POTENTIAL).
12   113   CLEAVAGE (BY SIMILARITY).
78   218   POTENTIAL.
40   240   N-LINKED (GLCNAC...) (POTENTIAL).
1 AA;  29360 MW;  10CA588D923754EB CRC64;

arity     7.5%;  Score 87;  DB 1;  Length 261;
onservative  25.9%;  Pred. No.1.8;   Indels  24;  Gaps  6;
21;  Mismatches  55;

WEE---TKINSSSPLRYDRQIGEFTVIRAGLYYLYCCVHFDEGKAVYLKLDLLVN 160
WAAEKGYYTMSNNLVTLENGKQL---TVKRQGLYYIYAQVTFCSNREASSQAPFIAS 192

LRCLEEF-----SATAASSPGPQLRLC-----QVSGLLPLRPGSSLRIRTLPWAH 209
LKPPNRFERILLRAANTHSSAKP----CGQQSIHLGGIFELQPGASVFNVTDPSQ 246

PFLTYFGLFQV 224
TGFTSFGLLKL 261

STANDARD;   PRT;   240 AA.
6; Q8WVF8; Q96LD2;
Rel. 40, (Created)
Rel. 40, Last sequence update)
Rel. 42, Last annotation update)
s factor ligand superfamily member 14 (Herpesvirus entry
nd) (HVEM-L).
GHT OR HVEML.
(Human).
tazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
heria; Primates; Catarrhini; Hominidae; Homo.
06;
1 N.A. (ISOFORM 1).
340; PubMed=9462508;
bner R., Montgomery R.I., Kochel K.D., Cheung T.C.,
en S., Murphy M., Eisenberg R.J., Cohen G.H., Spear P.G.,
, member of the TNF superfamily, and lymphotoxin alpha are
herpesvirus entry mediator.";
-30(1998).
1 N.A. (ISOFORM 1), AND CHARACTERIZATION.
532; PubMed=9765287;
McDonnell P.C., Brigham-Burke M., Lyn S.D., Minton J.,
le K., Spampanato J., Silverman C., Hensley P.,
Emery J.G., Deen K., Eichman C., Chabot-Fletcher M.,
oung P.R.;
entry mediator ligand (HVEM-L), a novel ligand for
```

; AAH18058.1; ALT_FRAME.
4TSV.
930; TNFSF14.

; F:receptor binding; TAS.
; P:induction of apoptosis; TAS.
; P:signal transduction; TAS.
06053; TNF_abc.
06052; TNF_family.
08983; TNF_like.
03636; TNF_subf.
     TNF; 1.
4; TNECROSISFCT.
012; TNF_subf; 1.
; TNF; 1.
251; TNF_1; FALSE_NEG.
049; TNF_2; 1.
nsmembrane; Glycoprotein; Signal-anchor;
licing.

```
1        240       TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
                   MEMBER 14, MEMBRANE FORM.
3        240       TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
                   MEMBER 14, SOLUBLE FORM.
1         37       CYTOPLASMIC (POTENTIAL).
8         58       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
                   (POTENTIAL).
9        240       EXTRACELLULAR (POTENTIAL).
2         83       CLEAVAGE (POTENTIAL).
2        187       POTENTIAL.
2        102       N-LINKED (GLCNAC.    .).
8         73       Missing (in isoform 2).
                   /FTId=VSP_006452.
0        120       L -> V (IN REF. 4).
4        214       E -> K (IN REF. 2).
. AA;   26351 MW;  49D0BF67E1390B39 CRC64;
```

```
urity    7.4%;  Score 86.5;  DB 1;  Length 240;
nservative 22.2%;  Pred. No. 1.8;
          30; Mismatches  87; Indels  83; Gaps  10;

LACLGLLLVVVSLG---SWATLSAQEPSQEELTAEDRREPPELNPQTEESQDVV 58
||||||||:::
RVGLGLLLLMGAGLAVQGWFLLQLHWRLGEMVT----RLP----------DGPA 77

LVRPRRSAPKGRKARPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSSS 118
|:: |||                   ::|||   :|::
LIQERRS----------------HEVNPAAHLTGANSSLTG--------SGG 110

RQI---------GEFTVIRAGLYLYLYCQVHPDE-----GKAVYL-------- 153
            ||   ||  |  |     | |
TQIGLAFLRGLSYHDGALVVTKAGYYYIYSKVQLGGVGCPLGLASTITHGLYXRT 170

KLDLLVNGVLALRCLEEFSATAASSPGPQLRLCQVSGLLPLRPGSSLRIRTLPWA 208
:||: |||        ||      |        |       :|||
ELELLVS------QQSPCGRATSSSRVWWDSSFLGGVVHLEAGEEVVVRVLDER 223

P-FLTYFGLFQV 224
  :||.||
DGTRSYFGAFMV 240
```

STANDARD;   PRT;   1237 AA.
; Q9ES10; Q9ES11; Q9ES12; Q9ES13;
el. 13, Created)
el. 13, Last sequence update)
el. 43, Last annotation update)
protein 2 (Non-erythroid band 3-like protein) (B3RP).
Mouse).
azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
eria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM A).
RX   MEDLINE=89034212; PubMed=3182834;
RA   Alper S.L., Kopito R.R., Libresco S.M., Lodish H.F.;
RT   "Cloning and characterization of a murine band 3-related cDNA fro
RT   kidney and from a lymphoid cell line.";
RL   J. Biol. Chem. 263:17092-17099(1988).
RN   [2]
RP   SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
RX   MEDLINE=20462926; PubMed=11006093;
RA   Lecanda J., Urtasun R., Medina J.F.;
RT   "Molecular cloning and genomic organization of the mouse AE2 anio
RT   exchanger gene.";
RL   Biochem. Biophys. Res. Commun. 276:117-124(2000).
CC   -!- FUNCTION: Plasma membrane anion exchange protein of wide
CC       distribution.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=5;
CC       Name=A;
CC         IsoId=P13808-1; Sequence=Displayed;
CC       Name=B1;
CC         IsoId=P13808-2; Sequence=VSP_000458;
CC       Name=B2;
CC         IsoId=P13808-3; Sequence=VSP_000457;
CC       Name=C1;
CC         IsoId=P13808-4; Sequence=VSP_000460;
CC       Name=C2;
CC         IsoId=P13808-5; Sequence=VSP_000459, VSP_000461;
CC   -!- TISSUE SPECIFICITY: Isoform a is widely expressed at similar
CC       levels in all tissues examined. Isoforms B1 and B2 are
CC       predominantly expressed in stomach although they are also dete
CC       at lower levels in other tissues. Isoform C1 is stomach-speci
CC       Isoform C2 is expressed at slightly higher levels in lung and
CC       stomach than in other tissues
CC   -!- SIMILARITY: Belongs to the anion exchanger family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a coll
CC   between the Swiss Institute of Bioinformatics and the EMBL out
CC   the European Bioinformatics Institute. There are no restriction
CC   use by non-profit institutions as long as its content is in
CC   modified and this statement is not removed. Usage by and for co
CC   entities requires a license agreement (See http://www.isb-sib.ch/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; J04036; AAA65505.1; -.
DR   EMBL; AF255774; AAG23154.1; -.
DR   EMBL; AF255774; AAG23155.1; -.
DR   EMBL; AF255774; AAG23156.1; -.
DR   EMBL; AF255774; AAG23158.1; -.
DR   EMBL; AF255774; AAG23157.1; -.
DR   PIR; A31789; A31789.
DR   HSSP; P02730; 1BTQ.
DR   MGD; MGI:109351; Slc4a2.
DR   InterPro; IPR001717; Anion_exchange.
DR   InterPro; IPR003020; HCO3_cotranspt.
DR   Pfam; PF00955; HCO3_cotransp; 1.
DR   PRINTS; PR01231; HCO3TRNSPORT.
DR   TIGRFAMs; TIGR00834; ae; 1.
DR   PROSITE; PS00219; ANION_EXCHANGER_1; 1.
DR   PROSITE; PS00220; ANION_EXCHANGER_2; 1.
KW   Transmembrane; Glycoprotein; Transport; Antiport; Ion transport;
KW   Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
FT   DOMAIN        1    703       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      704   1237       MEMBRANE (ANION EXCHANGER).
FT   TRANSMEM    704    727       POTENTIAL.
FT   TRANSMEM    733    770       POTENTIAL.
FT   TRANSMEM    790    812       POTENTIAL.
FT   TRANSMEM    822    843       POTENTIAL.
FT   DOMAIN      844    896       EXOPLASMIC LOOP (POTENTIAL).
FT   TRANSMEM    897    914       POTENTIAL.
FT   DOMAIN      915    929       CYTOPLASMIC (POTENTIAL).
```

```
30    950    POTENTIAL.
34   1006    POTENTIAL.
32   1053    POTENTIAL.
87   1132    POTENTIAL.
59   1195  .. POTENTIAL.
 5    316    PRO-RICH.
73     87    HIS-RICH.
61    865    POLY-SER.
55    855    N-LINKED (GLCNAC. . .)  (POTENTIAL).
66    866    N-LINKED (GLCNAC. . .)  (POTENTIAL).
78    878    N-LINKED (GLCNAC. . .)  (POTENTIAL).
69   1169    S-palmitoyl cysteine  (By similarity).
 1     17    MSSAPRRPASGADSLHT -> MDFLLRPQ (in isoform
             B2).
             /FTId=VSP_000457.
 1     17    MSSAPRRPASGADSLHT -> MTQ (in isoform B1).
             /FTId=VSP_000458.
 1    166    Missing (in isoform C2).
             /FTId=VSP_000459.
 1    198    Missing (in isoform C1).
             /FTId=VSP_000460.
67    193    ERTSPSPPTQTPHQEAAPRASKGAQTG -> MPAFQEWKSG
             GLREEAVFGAHGCSVCR (in isoform C2).
             /FTId=VSP_000461.
05    205    A -> G (IN REF. 2).
37 AA;  136813 MW;  1A0782C0071782EE CRC64;
```

```
arity    7.4%;  Score 86.5;  DB 1;  Length 1237;
onservative 27.9%;  Pred. No. 13;  Indels 77;  Mismatches 51;  Gaps 12;
19;  Mismatches 19;
```

```
TLSAQEPSQEELT-AEDRREPPELNPQTEESQDVVPFLEQ-----LVRPRRSAPKG 74
|: :| :|| | :| ||| :|  | :| |   | : |        : |:|| :||:
TIEEGEDEEEASEAEGFRAPPQ-QPSPATTPSAVQFFLQEDEGAERKPERTSPSP 173

PRRAIAAHYEVHPRPGQDGQAGQV------------------DGTVSGWEETKI---N 115
|  |||  |                                 :  |  || ::|
P-----HQEAAPRASK-GAQTGLVEEMVAVASATAGGDDGGAAGRPLTKAQPGH 226

LRYDRQIGEFTVIRAGLYYLYCQVHFDEGKAVL--KLDLL------VNGVL-- 163
|: :| :|:  | | | |:|:| ||  ||: |  :| || ||     ||
LQERRRIGSMTGVEQA--LLPRVPTDESEAQTLATADLDLMKSHRFEDVPGVRRH 283

LEEFSATRA---SSPGPQLR 184
|| | | |    ||||  |
CNAKGSTQAAREGREPGPTPR 307
```

```
STANDARD;    PRT;    261 AA.

(Rel. 25, Created)
(Rel. 25, Last sequence update)
(Rel. 43, Last annotation update)
.B factor ligand superfamily member 5 (CD40 ligand) (CD40-
ed activation protein) (TRAP) (T cell antigen Gp39)
n).
HOLG OR CD40L OR TRAP.
(Human)
etazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
heria; Primates; Catarrhini; Hominidae; Homo.
506;
```

```
N.A.
6854; PubMed=1280226;
chaeuer U., Mages H.W., Senger G., Kroczek R.A.;
TRAP, a ligand for CD40 on human T cells.";
mol. 22:3191-3194(1992).

N.A.
9181; PubMed=1385114;
D., Grosmaire L.S., Kullas C.D., Chalupny J.N.,
```

RA   Braesch-Andersen S., Noelle R.J., Stamenkovic I., Ledbetter J.A.,
RA   Aruffo A.;
RT   "The human T cell antigen gp39, a member of the TNF gene family,
RT   ligand for the CD40 receptor: expression of a soluble form of gp3
RT   with B cell co-stimulatory activity.";
RL   EMBO J. 11:4313-4321(1992).
RN   [3]
RP   SEQUENCE FROM N.A., AND VARIANTS HIGM1 128-ARG-GLY-129 AND PRO-23
RX   MEDLINE=93145330; PubMed=7678782;
RA   Aruffo A., Farrington M., Hollenbaugh D., Li X., Milatovich A.,
RA   Nonoyama S., Bajorath J., Grosmaire L.S., Stenkamp R., Neubauer M
RA   Roberts R.L., Noelle R.J., Ledbetter J.A., Francke U., Ochs H.D.;
RT   "The CD40 ligand, gp39, is defective in activated T cells from
RT   patients with X-linked hyper-IgM syndrome.";
RL   Cell 72:291-300(1993).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93094757; PubMed=1281209;
RA   Spriggs M.K., Armitage R.J., Strockbine L., Clifford K.N.,
RA   Macduff B.M., Sato T.A., Maliszewski C.R., Fanslow W.C.;
RT   "Recombinant human CD40 ligand stimulates B cell proliferation an
RT   immunoglobulin E secretion.";
RL   J. Exp. Med. 176:1543-1550(1992).
RN   [5]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93138085; PubMed=7678552;
RA   Gauchat J.F.M., Aubry J.-P., Mazzei G.J., Life P., Jomotte T.,
RA   Elson G., Bonnefoy J.Y.;
RT   "Human CD40-ligand: molecular cloning, cellular distribution and
RT   regulation of expression by factors controlling IgE production.";
RL   FEBS Lett. 315:259-266(1993).
RN   [6]
RP   SEQUENCE FROM N.A.
RA   Shimadzu M., Terasaki H., Ninomiya R., Shimizu S., Nunoi H.,
RA   Matsuda I.;
RL   Submitted (FEB-1995) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE OF 113-117, AND PROCESSING.
RX   MEDLINE=96198042; PubMed=8626375;
RA   Pietravalle F., Lecoanet-Henchoz S., Blasey H., Aubry J.-P., Els(
RA   Edgerton M.D., Bonnefoy J.-Y., Gauchat J.-F.;
RT   "Human native soluble CD40L is a biologically active trimer, pro(
RT   inside microsomes.";
RL   J. Biol. Chem. 271:5965-5967(1996).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 116-261.
RX   MEDLINE=96131874; PubMed=8589998;
RA   Karpsusas M., Hsu Y.-M., Wang J.-H., Thompson J., Lederman S.,
RA   Chess L., Thomas D.;
RT   "2-A crystal structure of an extracellular fragment of human CD4(
RT   ligand.";
RL   Structure 3:1031-1039(1995).
RN   [9]
RP   3D-STRUCTURE MODELING OF COMPLEX WITH CD40.
RX   MEDLINE=98266353; PubMed=9605317;
RA   Singh J., Garber E., van Vlijmen H., Karpsusas M., Hsu Y.-M.,
RA   Zheng Z., Naismith J.H., Thomas D.;
RT   "The role of polar interactions in the molecular recognition of (
RT   with its receptor CD40.";
RL   Protein Sci. 7:1124-1135(1998).
RN   [10]
RP   VARIANTS HIGM1 ARG-36 AND GLY-140.
RX   MEDLINE=93156839; PubMed=7679206;
RA   Korthaeuer U., Graf D., Mages H.W., Briere F., Padayachee M.,
RA   Malcolm S., Ugazio A.G., Notarangelo L.D., Levinsky R.J.,
RA   Kroczek R.A.;
RT   "Defective expression of T-cell CD40 ligand causes X-linked
RT   immunodeficiency with hyper-IgM.";
RL   Nature 361:533-541(1993).
RN   [11]
RP   VARIANT HIGM1 GLU-123.
RX   MEDLINE=93156840; PubMed=8094231;
RA   Disanto J.P., Bonnefoy J.Y., Gauchat J.F.M., Fischer A.,

e G.;
utations in X-linked immunodeficiency with hyper-IgM.";
-543 (1993).

PRO-155; ASP-211 AND VAL-227.
70; PubMed=769801;
mitage R.J., Conley M.E., Rosenblatt H., Jenkins N.A.,
Bedell M.A., Edelhoff S., Disteche C.M.,
, Fanslow W.C., Belmont J.W., Spriggs M.K.;
ene defects responsible for X-linked hyper-IgM
0-993 (1993).

ALA-126, ARG-140 AND GLU-144.
38; PubMed=7717401;
la A., Strina D., Sacco M.G., Morali F., Brugnoni D.,
ntuano E., Fasth A., Andersson B., Zegers B.J.M.,
znick I., Levy J., Zan-Bar I., Porat Y., Airo P.,
zzoni P., Notarangelo L.D.;
ion of nine novel mutations in the CD40 ligand gene in
X-linked hyper IgM syndrome of various ancestry.";
net. 56:898-906(1995).

PRO-155 AND VAL-227, AND VARIANT ARG-219.
33; PubMed=8550833;
J., Allen R.C., Larche M., Greene J.M., Shigeoka A.O.,
rauf D.C., Belmont J.W., Conley M.E.;
nd conformation polymorphism study of CD40 ligand.
tion analysis and carrier detection for X-linked hyper
t. 97:196-201(1996).

ARG-36; CYS-140; SER-231; MET-254 AND GLY-227 DEL.
77; PubMed=9150729;
himadzu M., Toru H., Seyama K., Nunoi H., Neubauer M.,
h H.D.;
the CD40 ligand gene in 13 Japanese patients with
-IgM syndrome.";
:624-627(1997).

Mediates B-cell proliferation in the absence of co-
s well as IgE production in the presence of IL-4.
n immunoglobulin class switching.
omotrimer.
R LOCATION: Type II membrane protein. Also exists as an
lar soluble form.
CIFICITY: Specifically expressed on activated CD4+
tes.
oluble form derives from the membrane form by
c processing.
efects in TNFSF5 are the cause of X-linked
ciency with hyper-IgM type 1 (HIGM1) [MIM:308230]. HIGM1
noglobulin isotype switch defect characterized by
oncentrations of serum IgM and decreased amounts of all
ypes. Affected males present at an early age (usually
first year of life) recurrent bacterial and
tic infections, including pneumocystis carinii pneumonia
table diarrhea due to cryptosporidium infection. Despite
on treatment with intravenous immunoglobulin, the
ognosis is rather poor, with a death rate of about 10%
olescence.
: Belongs to the tumor necrosis factor family.
NAME=CD40lbase;
ean CD40L defect database (mutation db).
//www.expasy.org/cd40lbase/";
//ftp.expasy.org/databases/cd40lbase/
NAME=PROW; NOTE=CD guide CD154 entry;
//www.ncbi.nlm.nih.gov/prow/cd/cd154.htm".

DR   EMBL; X68550; CAA48554.1; -.
DR   EMBL; Z15017; CAA78737.1; -.
DR   EMBL; X67878; CAA48077.1; -.
DR   EMBL; L07414; AAA35662.1; -.
DR   EMBL; D31797; BAA06599.1; JOINED.
DR   EMBL; D31793; BAA06599.1; JOINED.
DR   EMBL; D31794; BAA06599.1; JOINED.
DR   EMBL; D31795; BAA06599.1; JOINED.
DR   EMBL; D31796; BAA06599.1; JOINED.
DR   PIR; S28017; I53476.
DR   PDB; 1ALY; 17-SEP-97.
DR   PDB; 1I9R; 22-MAY-02.
DR   Genew; HGNC:11935; TNFSF5.
DR   MIM; 300386; -.
DR   MIM; 308230; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0005625; C:soluble fraction; TAS.
DR   GO; GO:0005174; F:CD40 receptor binding; IPI.
DR   GO; GO:0006916; P:anti-apoptosis; IDA.
DR   GO; GO:0042100; P:B-cell proliferation; IDA.
DR   GO; GO:0006954; P:inflammatory response; IDA.
DR   GO; GO:0045190; P:isotype switching; ISS.
DR   GO; GO:0007159; P:leukocyte cell adhesion; NAS.
DR   GO; GO:0030168; P:platelet activation; IDA.
DR   GO; GO:0007165; P:signal transduction; ISS.
DR   InterPro; IPR003263; TNF_5.
DR   InterPro; IPR006052; TNF_family.
DR   InterPro; IPR008983; TNF_like.

```
Query Match             7.4%;   Score 86;   DB 1;   Length 261;
Best Local Similarity  25.9%;   Pred. No. 2.2;
Matches  35;  Conservative  21;  Mismatches  55;  Indels  24;  Ga

Qy  105 TVSGWEE---TKINSSSPLRYDRQIGEFTVIRAGLYYLYCQVHFDEGKAVYLKLDI
            :|.:|   ||:.:|  ||::||:|.: |::|:|   :
Db  136 SVLQWAEKGYYTMSNNLVTLENGKQL---TVKRQGLYYIYAQVTFCSNRBASSQAPI

Qy  161 GVLALRCLEEF------SATAASPGPQLRLC-----QVSGLLPLRPGSSLRIRTLI
            :|  |:|:|   |:  |:.||::|   ||:|:|
Db  193 --LCLKSPGRFERILLRAANTHSSAKP----CGQQSIHLGGVFELQPGASVFVNVTI

Qy  210 LKAAPFLTYFGLFQV 224
            ||||  ::
Db  247 VSHGTGFTSFGLLKL 261
```

Search completed: April 7, 2004, 17:45:17
Job time : 10.2829 secs

GenCore version 5.1.6
yright (c) 1993 - 2004 Compugen Ltd.

search, using sw model

7, 2004, 17:41:27 ; Search time 31.3851 Seconds
(without alignments)
2261.954 Million cell updates/sec

)-245-198a-2

SLGLALACLGLLLVVVSL.........PWAHLKAAPFLTYFGLFQVH 225

JM62
10.0 , Gapext 0.5

041 seqs, 315518202 residues

satisfying chosen parameters: 1017041

1: 0
1: 2000000000

imum Match 0%
imum Match 100%
ing first 45 summaries

REMBL_25:*
sp_archea:*
sp_bacteria:*
sp_fungi:*
sp_human:*
sp_invertebrate:*
sp_mammal:*
sp_mhc:*
sp_organelle:*
sp_phage:*
sp_plant:*
sp_rodent:*
sp_virus:*
sp_vertebrate:*
sp_unclassified:*
sp_rvirus:*
sp_bacteriap:*
sp_archeap:*

he number of results predicted by chance to have a
than or equal to the score of the result being printed,
by analysis of the total score distribution.

SUMMARIES

Y	Length	DB	ID	Description
1	410	11	Q8BXS2	Q8bxs2 mus musculu
8	330	4	Q8IZK7	Q8izk7 homo sapien
0	409	5	Q8IGD3	Q8igd3 drosophila
5	409	5	Q8MY88	Q8my88 drosophila
3	261	5	Q8MRW2	Q8mrw2 drosophila
3	325	5	Q9V5G2	Q9v5g2 drosophila
3	415	5	Q8MUJ1	Q8muj1 drosophila
9	426	16	Q88IZ6	Q88iz6 pseudomonas
4	684	13	Q7T3J6	Q7t3j6 lampetra ja
0	557	16	Q8XQX3	Q8xqx3 ralstonia s
8	210	16	Q9A926	Q9a926 caulobacter
8	421	16	Q9HUW2	Q9huw2 pseudomonas
7	213	16	Q82AD2	Q82ad2 streptomyce
7	1363	5	Q9VFD3	Q9vfd3 drosophila
7	287	13	Q90WT9	Q90wt9 gallus gall
6	224	5	Q9V762	Q9v762 drosophila

17	88	7.6	352	12	O89341	O89341 hendr;
18	88	7.6	353	12	Q66760	Q66760 equin;
19	88	7.6	532	4	Q16727	Q16727 homo s;
20	87.5	7.5	522	10	Q9FTN7	Q9ftn7 oryza
21	87.5	7.5	670	6	Q9AA15	Q9aa15 caulo;
22	87	7.5	244	6	Q862Z7	Q862z7 pan tr;
23	86.5	7.4	154	16	Q7U7N8	Q7u7n8 synec;
24	86.5	7.4	340	16	Q9HUR8	Q9hur8 pseud;
25	86.5	7.4	504	16	Q92KA4	Q92ka4 rhizo;
26	86.5	7.4	1237	11	Q7TPS4	Q7tps4 mus m;
27	86	7.4	260	10	Q8S2N9	Q8s2n9 oryza
28	85.5	7.4	331	10	Q942P9	Q942p9 oryza
29	85.5	7.4	724	5	Q868S9	Q868s9 anophe;
30	85.5	7.4	2841	16	Q7XU06	Q7xu06 oryza
31	85	7.3	116	16	Q7V2A2	Q7v2a2 proch;
32	84.5	7.3	377	16	Q7UEA6	Q7uea6 rhodo;
33	84.5	7.3	422	16	Q9RKB0	Q9rkb0 strept;
34	84.5	7.3	430	2	Q9REU1	Q9reu1 strept;
35	84.5	7.3	1079	13	Q8UVR4	Q8uvr4 xenop;
36	84.5	7.3	1118	16	Q98E34	Q98e34 rhizo;
37	84	7.2	2962	5	Q93326	Q93326 caenor;
38	83.5	7.2	394	16	Q92V66	Q92v66 rhizo;
39	83	7.1	467	16	Q9S2Y4	Q9s2y4 strep;
40	82.5	7.1	174	16	Q9CKX1	Q9ckx1 paste;
41	82.5	7.1	314	5	Q8WPH7	Q8wph7 theile;
42	82.5	7.1	501	16	Q89IN3	Q89in3 brady;
43	82.5	7.1	549	16	Q8RC38	Q8rc38 therm;
44	82.5	7.1	718	6	Q8HXH0	Q8hxh0 macaca
45	82	7.1	619	5	Q8SUH9	Q8suh9 enceph;

ALIGNMENTS

RESULT 1
Q8BXS2
ID   Q8BXS2   PRELIMINARY;   PRT;   410 AA.
AC   Q8BXS2;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Tumor necrosis factor.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Retina;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK044387; BAC31897.1; -.
DR   PIR; PT0714; PT0714.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005164; F:tumor necrosis factor receptor binding; IEA.
DR   GO; GO:0006955; P:immune response; IEA.
DR   InterPro; IPR006052; TNF family.
DR   InterPro; IPR008983; TNF_like.
DR   SMART; SM00207; TNF_2.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS55049; TNF_2; 2.
SQ   SEQUENCE 410 AA; 45881 MW; 590A4B74C33FB8D4 CRC64;

Query Match   82.1%;   Score 954.5;   DB 11;   Length 410;
Best Local Similarity   88.6%;   Pred. No. 1.3e-83;
Matches 195;   Conservative 1;   Mismatches 13;   Indels 11;   Ga)

Qy   1 VLSLGLALACLGLLLVVVSLGSWATLSA-QEPSQELTAEDRREPPELNPQTEESQD'

```
GN   BCDNA:RH51659.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=y;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Dresnek D., Parfan D., Frise E.,
RA   George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA   Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S
RA   Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.
RA   Celniker S.;
RL   Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BT001838; AAN71595.1; -.
DR   FlyBase; FBgn0064801; BcDNA:RH51659.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005164; F:tumor necrosis factor receptor binding; IEA.
DR   GO; GO:0006955; P:immune response; IEA.
DR   InterPro; IPR006052; TNF_family.
DR   InterPro; IPR008983; TNF_like.
DR   SMART; SM00207; TNF; 1.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 1.
SQ   SEQUENCE   409 AA;  46401 MW;  FC2E9BD9BE012D257 CRC64;

Query Match             10.0%;  Score 116.5;  DB 5;  Length 409;
Best Local Similarity   24.3%;  Pred. No. 0.01;
Matches   50;  Conservative   32;  Mismatches   87;  Indels   37;  G

Qy    29 QEPSQEELTAEDRREPPELNPQTEESQDVVPFLEQLVRPRRSAPKGRKARPRRAIA
         ||::||:|   :::||  |          :|:||:|
Db   231 QEKSSNEATSKERPAPLHHRRRMHSRH----RHLLVRKARS----EDSRP----A

Qy    89 VHPRPGQDGAQAGVDGTVSGWEETKINSSSPLRYDRQIGEFTVIRAGLYYLYCQVH
         :  |:: |  |  | ||| ::::     ||:| ||:|:|  || |||:|| |||
Db   278 LSSRRHQQSM-GYHGDMYIGNDNERNSYQG-HFQTRDGVLTVNTGLYYVYAQIC

Qy   149 KAVVLKDLLVNGVLA------LRCLEEFSATAASSPGPLRLCQVSGLLPLRPG
         :  :  |:|:||:       |||||:|:| ||  ::||   ::| |||
Db   336 HD--------QNGFIVFQGDTPFLQCLN---TVPTNMPHKVHTCHTSGLIHLERNE

Qy   202 IRTL--PWAHLKAAPFLTYFGLFQV 224
         :|:   ||||:||||:|:|||||:|
Db   384 LKDIHNDRNAVLREGNNRSYFGIFKV 409

RESULT 4
Q8MY88
ID   Q8MY88
AC   Q8MY88;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   TNF superfamily ligand, Eiger (Tumor necrosis factor family member
DE   DT1).
GN   EIGER OR DT1 OR CG12919.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22060500; PubMed=12065414;
RA   Igaki T., Kanda H., Yamamoto-Goto Y., Kanuka H., Kuranaga E.,
RA   Aigaki T., Miura M.;
RT   "Eiger, a TNF superfamily ligand that triggers the Drosophila JNK
RT   pathway.";
RL   EMBO J. 21:3009-3018(2002).
RN   [2]
```

```
GLALACLGLLLVVSLGSWATLSAQQEPSQEELTAEDRREPPELNPQTEESQDVVP 84

LVRPRRSAPKGRKARPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSSSP 119
||||||||||||||||||||||||||||||||||||||||||||||||||||||
LVRPRRSAPKGRKARPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSSSP 144

RQIGEFTVIRAGLYYLYCQVHFDEGKAVYLKDLLVNGVLALRCLEEFSATAASSP 179
||||||||||||||||||||||||||||||||||||||||||||||||||||||
RQIGEFTVIRAGLYYLYCQVHFDEGKAVYLKDLLVNGVLALRCLEEFSATAASSP 204

RLCQVSGLLPLR--------PGSSLRIRTLPWAHL 210
|| |  ||  |    ||  |
RLCQTE-LQSLRREVSRLQRSGGPSQKQGGERPWQSL 243
```

```
RELIMINARY;   PRT;   330 AA.

TrEMBLrel. 23, Created)
TrEMBLrel. 23, Last sequence update)
TrEMBLrel. 25, Last annotation update)
```

```
(Human).
azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
eria; Primates; Catarrhini; Hominidae; Homo.
06;.

N.A.
24; PubMed=12411489;
B., Medema J.P., Lopez-Fraga M., Lozano J.C.,
M., Picard A., Martinez-A C., Garcia-Sanz J.A.,

  hybrid mRNA encodes TWE-PRIL, a functional cell surface
usion protein.";
1-5720(2002).
; AAL90443.1; -.
; C:membrane; IEA.
; F:tumor necrosis factor receptor binding; IEA.
; P:immune response; IEA.
06052; TNF_family.
08983; TNF_like.
 TNF; 1.
; TNF; 1.
51; TNF 1; 1.
49; TNF 2; 2.
AA;  36588 MW;  FC6F3BCA29C029AE CRC64;

         52.8%;  Score 613;  DB 4;  Length 330;
rity     84.4%;  Pred. No. 9e-51;
nservative   8;  Mismatches   14;  Indels   0;  Gaps   0;

ALACLGLLLVVSLGSWATLSAQEPSQEELTAEDRREPPELNPQTEESQDVVPFL 61
||||||||||||||||:||: :::|||||||||||||||||||||||||||:||
ALACLGLLLAVVSLGSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQDPAPFL 85

PRRSAPKGRKARPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSSSPLR 121
||||||||||:| |||||||||||||||||||||||||||||||| :|||||||
PRRSAPKGRKTRARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLR 145

GEFTVIRAGLYYLYCQ 142
|||:| |||||||||
GEFIVTRAGLYYLYCQ 166
```

```
ELIMINARY;   PRT;   409 AA.

EMBLrel. 23, Created)
EMBLrel. 23, Last sequence update)
EMBLrel. 25, Last annotation update)
```

Query Match   9.3%;   Score 107.5;   DB 5;   Length 261;
Best Local Similarity   23.1%;   Pred. No. 0.042;   --
Matches   48; Conservative   34; Mismatches   91; Indels   35; Gap

Qy   29 QEPSQEELTAEDRREPPELNPQTEESQDVVPFLEQLVRPRRS--APKGRKARPRRAIA
          :||: ::: : ::||| |:|| |||
Db   77 QEKSSNEATSKESPAPLHHRRMHSRH----RHLLVRKGESLLSARSEDSRP----A

Qy   87 YEVHPRPGQDGAQAGVDGTVSGWEETKINSSSPLRYDRQIGEFTVIRAGLIYYLYCQVH
          : || | ||| || | ||| ||: : :::| ||||: |: |||||
Db   128 FHLSSRRRHQGSM-GYHGDMYIGNDNERNSYQG-HFQTRDGVLTVTNTGLYYYAQIC

Qy   147 EGKAVYLKLDLLVNGVLA------LRCLEEFSATAASSPGPQLRLCQVSGLLPLRPG
                           ||| ||:|
Db   186 NSHD-------QNGFIVFQGDTPFLQCLN----TVPTNMPHKVHTCHTSGLIHLERN

Qy   200 LRIRTL--PWAHLKAAPFLTYFGLFQV 224
                 |||:||::|:|
Db   234 IHLKDIHNDRNAVLREGNNRSYFGIFKV 261

RESULT 6
Q9V5G2
ID   Q9V5G2   PRELIMINARY;   PRT;   325 AA.
AC   Q9V5G2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   CG12919 protein.
GN   EIGER OR CG12919.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;

RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra J
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L

DR   InterPro; IPR006052; TNF family.
DR   InterPro; IPR008983; TNF_like.
DR   SMART; SM00207; TNF; 1.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 1.
SQ   SEQUENCE   415 AA; 46918 MW; E087A26DE222D2BF CRC64;

Query Match   9.3%; Score 107.5; DB 5; Length 415;
Best Local Similarity   23.1%; Pred. No. 0.077;
Matches   48; Conservative   34; Mismatches   91; Indels   35; C

Qy   29 QEPSQELTAEDRREPPELNPQTEESQDVVPFLEQLVRPRRS--APKGRKARPRRA
              ||| |||||| : : :|| : : :: | :| ||: :||
Db   231 QEKSSNEATSKESPAPLHHRRRMHSRH----RHLLVRKGESLLSARSEDSRP--

Qy   87 YEVHPRPGQDGAQAGVDGTVSGWEETKINSSSPLRYDRQIGEFTVIRAGLYYLYCQ
             | | : : ||| : : || : | || |:: :|||:|:|
Db   282 FHLSSRRHQGSM-GYHGDMYIGNDNERNSYQG-HFQTRDGVLTVNTGLYYVYAQ

Qy   147 EGKAVYLKLDLLVNGVLA-----LRCLEEFSATAASSPGPQLRLCQVSGLLPLR
                   : || :|| : :   ::| :|:: :
Db   340 NSHD-------QNGFIVFQGDTPFIQCLN---TVPTNMPHKVHTCHTSGLIHLE

Qy   200 LRIRTL--PWAHLKAAPFLTYFGLFQV 224
                 : | ::|| :| :|
Db   388 IHLKDIHNDRNAVLREGNNRSYFGIFKV 415

RESULT 8
Q88IZ6
ID   Q88IZ6   PRELIMINARY;   PRT;   426 AA.
AC   Q88IZ6;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Conserved hypothetical protein.
GN   PP2853.
OS   Pseudomonas putida (strain KT2440).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=160488;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22431060; PubMed=12534463;
RA   Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA   Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes t
RA   Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
RA   Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
RA   Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzez A.,
RA   Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
RA   Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA   Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA   Fraser C.M.;
RT   "Complete genome sequence and comparative analysis of the
RT   metabolically versatile Pseudomonas putida KT2440.";
RL   Environ. Microbiol. 4:799-808(2002).
DR   EMBL; AE016784; AAN68461.1; -.
DR   TIGR; PP2853; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   426 AA; 46020 MW; FEDC7E266C982633 CRC64;

Query Match   8.9%; Score 103; DB 16; Length 426;
Best Local Similarity   29.1%; Pred. No. 0.22;
Matches   50; Conservative   17; Mismatches   65; Indels   40; Ga

Qy   7 ALACLGLLLVVVSLGSWATLSAQEPSQELTAEDRREPPELNPQTEESQDVVPFLEC
              ||| | ||||| :| |: |||||||||||||||||||||||||
Db   9 AFVCLTTLAPATAL----YAAADPQVEAL----RQELIELKRRYEAQQQALMVLEC

Qy   67 PRRSAPKGRKARPRRAIAAHYEVHPRPGQDGAQ--AGVDGTV-SGWEETKINSSSP
              | :||| | :| :| |||:|| :| ::|: |: : |
Db   60 QVEEAPA--AQPKRLVKS-----PAEGVKGAQTVASGAPGTTGSSYGQALTADSEP

RA   hong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   yers E.W., Rubin G.M., Venter J.C.;
RL   equence of Drosophila melanogaster.";
RL   185-2195(2000).
DR   1; AAF58848.1; -.
DR   0033483; eiger.
DR   0; C:membrane; IEA.
DR   4; F:tumor necrosis factor receptor binding; IEA.
DR   5; P:immune response; IEA.
DR   006052; TNF_family.
DR   008983; TNF_like.
DR   7; TNF; 1.
DR   251; TNF_1; 1.
DR   049; TNF_2; 1.
SQ   5 AA; 36862 MW; 6E5CCB69694F1A3A CRC64;

arity   9.3%; Score 107.5; DB 5; Length 325;
                23.1%; Pred. No. 0.056;
onservative   34; Mismatches   91; Indels   35; Gaps   9;

QEELTAEDRREPPELNPQTEESQDVVPFLEQLVRPRRS--APKGRKARPRRAIAAH 86
   ||| |||||| : : :|| : : :: | :| ||: :||| |||
SNEATSKESPAPLHHRRRMHSRH----RHLLVRKGESLLSARSEDSRP---AAH 191

PRPGQDGAQAGVDGTVSGWEETKINSSSPLRYDRQIGEFTVIRAGLYYLYCQVHFD 146
   : : ||| : : || : | || |:: :|||:|:| : :
SRRRHQGSM-GYHGDMYIGNDNERNSYQG-HFQTRDGVLTVNTGLYYVYAQICYN 249

VYLKLDLLVNGVLA-----LRCLEEFSATAASSPGPQLRLCQVSGLLPLRPGSS 199
         : || :|| : :   ::| :|:: : ||
------QNGFIVFQGDTPFIQCLN---TVPTNMPHKVHTCHTSGLIHLERNER 297

TL---PWAHLKAAPFLTYFGLFQV 224
         : | ::|| :| :|
DIHNDRNAVLREGNNRSYFGIFKV 325

ELIMINARY;   PRT;   415 AA.

rEMBLrel. 22, Created)
rEMBLrel. 22, Last sequence update)
rEMBLrel. 25, Last annotation update)

19 OR DARTH.
anogaster (Fruit fly).
azoa; Arthropoda; Hexapoda; Insecta; Pterygota;
pterygota; Diptera; Brachycera; Muscomorpha;
rosophilidae; Drosophila.
7;

N.A.
23; PubMed=12176339;
M., Basler K.;
TNF Signaling Mechanisms. JNK-Dependent Apoptosis
iger, the Drosophila Homolog of the TNF Superfamily.";
:1263-1268(2002).

N.A.
38; PubMed=12894227;
aaty W.S., Chen P., Tomar R.S., Eby M.T., Chapo J.,
re N., Zachariah S., Sinha S.K., Abrams J.M.,

receptor, Wengen, comprise a TNF-like system in

60-486712(2003).
; AAM76710.1; -.
; AAM66763.1; -.
033483; eiger.
; C:membrane; IEA.
; F:tumor necrosis factor receptor binding; IEA.
; P:immune response; IEA.

```
DR   EMBL; AL646082; CAD18246.1; -.
DR   GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW   Plasmid; Complete proteome.
SQ   SEQUENCE   557 AA;  60786 MW;  8B974C147D710649 CRC64;

Query Match             8.0%;  Score 93.5;  DB 16;  Length 557;
Best Local Similarity  26.6%;  Pred. No. 2.5;
Matches   62;  Conservative   25;  Mismatches   77;  Indels   69;  Ga

Qy     2 LSLGLALCLGLLLVVSLGSWATLS-AQEPSQELTAEDRREPPR--LNPQTEESQI
           |:|:||  |||||:||||||||  :||:::|| : : :  ||
Db    10 LATGVALA-------GPLSSWAQSSPAQGNPIDTLPRVDTSRPPEQKIHVQVQRPN-

Qy    59 PFLEQL----VRPRRSAPKGRKARPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEE
           |||| |  : ||||| |||| || ||  ||||||  |:|:|||||||||||||
Db    59 PALENLLATHLTPSKFQIEGVKALPPEIAAHFA--PMAGHDVTVAQLLQAAS--BV

Qy   115 --NSSSPLRYDRQIGEFTVIRAGLYLYCQVHFDEGKAVVLKLDLLVNGVLALRLCLE
              : : : :                        ||||||||||| :| | :||
Db   115 YADRGYPL------SFAFVPAQTF--------EGGVVRI---TVVEGYVARMRIE

Qy   173 ATAASSPGP---QLRLC---------------------QVSGLLPLRPGSSLRIRTLP 206
              |||        |||                     :|:|:| :||
Db   153 ---GKPGPLEARLRAISKHMMDERPLRRETFERVTGVLALQPGVQITATVQP 201

RESULT 11
Q9A926
ID   Q9A926           PRELIMINARY;         PRT;    210 AA.
AC   Q9A926;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein CC1168.
GN   CC1168.
OS   Caulobacter crescentus.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC   Caulobacteraceae; Caulobacter.
OX   NCBI_TaxID=155892;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 19089 / CB15;
RX   MEDLINE=21173698; PubMed=11259647;
RA   Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.
RA   Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.
RA   Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely
RA   DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA   Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.
RA   Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White
RA   Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT   "Complete genome sequence of Caulobacter crescentus.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR   EMBL; AE005794; AAK23152.1; -.
DR   PIR; D87394; D87394.
DR   TIGR; CC1168; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   210 AA;  22344 MW;  8B9830ADFBF7F45C CRC64;

Query Match             7.8%;  Score 90.5;  DB 16;  Length 210;
Best Local Similarity  26.2%;  Pred. No. 1.4;
Matches   33;  Conservative   19;  Mismatches   67;  Indels    7;  Ga

Qy     2 LSLGLALACLGLLLVVSLGSWATLSAQEPSQELTAEDRREPPELNPQTEESQDVV
           |||| |                :|||:::          ||:||
Db    32 LRLGLAVFAAGVAATVIVQAAWRSLASSKLQTQAATAPLVLDKPRF---TGVLKDGR

Qy    62 EQLVRPRRSAPKGRKARPRRAIAAHYEVHPRPGQDGAQAGV----DGTVSGWEETKI
           : ||: : |||||:|||||||||||||| | | || :|   |||||||
Db    89 ITAERAERDAKDQNIVRLTAPLLVRGYGEPNPSQATAKSGVYREAENTLLLTDEVKI

Qy   118 SPLRYD 123
         |||:||
Db   149 EGFDFD 154
```

Left side column:
```
RQIGEF---TVIRAGLYLYCQVHFDEGKAVVLKLDLLVNGVLAL 165
         |: |:|  || :: :       || | :||  |||
EASGFPGGGKFSFETGVTY---THYDTRA------LTLNGFLAL 153

LIMINARY;      PRT;    684 AA.

EMBLrel. 25, Created)
EMBLrel. 25, Last sequence update)
EMBLrel. 25, Last annotation update)
protein.
ca (Japanese lamprey) (Entosphenus japonicus).
zoa; Chordata; Craniata; Vertebrata; Hyperoartia;
mes; Petromyzontidae; Lethenteron.
9;

.A.

T.;
ing of C4bp/Cremp-like protein.";
2001) to the EMBL/GenBank/DDBJ databases.
BAC77070.1; -.
AA;  76964 MW;  124D0D555909DC26 CRC64;

                8.4%;  Score 98;  DB 13;  Length 684;
ity   25.5%;  Pred. No. 1.2;
servative   17;  Mismatches   57;  Indels   72;  Gaps   10;

SWATLSAQEPSQELTAEDRREPPELNPQTEESQDVVPFLEQ---LVRPRRS- 70
:  ||||  | :|:||      || ::::| :
T--TFQTSPPEQESLT-----PPQSIPPLFNTMDVAPNVQQTPIIIVPPRPSP 110

RKARPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSSSPLRYDRQIGE 128
   |                 :    VVSGWSEDIYREGHLVTYACQSG- 151
SCSTP--------PDLPHSD-----

GLYLYCQVHFDEGKAVVLKLDLLVNGVLALRLCLEEFSATAASSPGPQLR---- 184
: |:                                      :|
RRFSIVCESDND------------------------------SPG-KLRWSSP 179

SGLLPLRPGS 198
 |  |
TCGRPENPAN 195

LIMINARY;      PRT;    557 AA.

EMBLrel. 20, Created)
EMBLrel. 20, Last sequence update)
EMBLrel. 24, Last annotation update)
tion/secretion signal peptide protein.
601.
nacearum (Pseudomonas solanacearum).
smid.
obacteria; Betaproteobacteria; Burkholderiales;
e; Ralstonia.

N.A.

79; PubMed=11823852;
Genin S., Artiguenave F., Gouzy J., Mangenot S.,
ault A., Brottier P., Camus J.C., Cattolico L.,
oisne N., Claudel-Renard C., Cunnac S., Demange N.,
ie M., Moisan A., Robert C., Saurin W., Schiex T.,
ebault P., Whalen M., Wincker P., Levy M.,
Boucher C.A.;
e of the plant pathogen Ralstonia solanacearum.";
-502(2002).
```

RA   Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT   "Genome sequence of an industrial microorganism Streptomyces
RT   avermitilis: deducing the ability of producing secondary
RT   metabolites.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX   MEDLINE=22608306; PubMed=12692562;
RA   Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba
RA   Sakaki Y., Hattori M., Omura S.;
RT   "Complete genome sequence and comparative analysis of the industri
RT   microorganism Streptomyces avermitilis.";
RL   Nat. Biotechnol. 21:526-531(2003).
DR   EMBL; AP005045; BAC73838.1; -.
DR   InterPro; IPR007561; DUF552.
DR   Pfam; PF04472; DUF552; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   213 AA; 23926 MW; 826FD2DD2D910030 CRC64;

Query Match   7.7%;
Best Local Similarity   27.7%;   Pred. No. 1.6;
Matches   41; Conservative   14; Mismatches   45; Indels   48; G

Qy   30 EPSQEELTAEDRR--EPPE--------LNPQTEES----------QDVVPPLEQL----V

Db   32 EPELDPEPERDRRRHEPPHQSHQALHPQRDESVRVVQPPMQRDPVPHSASLPAESV

Qy   70 SAPKGRKARPRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSSSPLRYD--

Db   92 IAPVASITQERQSL---------EKNAPVIMPKVVSEREPYRITTLHPRTYNEA

Qy   128 EFTVIRAGLIYLYCQVHFDEGKAVYLKL 155

Db   141 E-------------------HFREGTPVIMNL 153

RESULT 14
Q9VFD3
ID   Q9VFD3   PRELIMINARY;   PRT;   1363 AA.
AC   Q9VFD3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   CG31304 protein.
GN   CG31304 OR CG6960 OR CG18442.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.I
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin I
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong Z., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

---

PRELIMINARY;   PRT;   421 AA.

TrEMBLrel. 16, Created)
TrEMBLrel. 16, Last sequence update)
TrEMBLrel. 24, Last annotation update)
protein PA4851.

.eruginosa.
oteobacteria; Gammaproteobacteria; Pseudomonadales;
eae; Pseudomonas.
7;

N.A.
5692 / PAO1;
337; PubMed=10984043;
Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
oulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
pencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
ier M.H., Hancock R.E.W., Lory S., Olson M.V.;
ome sequence of Pseudomonas aeruginosa PAO1, an
pathogen.";
9-964(2000).
8; AAG08236.1; -.
C83040.
protein; Complete proteome.
1 AA; 46442 MW; 68CSC17099953C3E CRC64;

arity   7.8%;   Score 90.5; DB 16;   Length 421;
onservative   27; Mismatches   76; Indels   79; Gaps   10;

QEELTAEDR--------REPPELNPQTEESQDVVPFL-------------EQLV 65
           |||||| ::   ::
GEPLAEERVSFVPLRRAEDPPELRLSLDDDDDEPSLGKRRHGLSALDEDDEESA 222

SAPKGRKARPRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSSSPLRYDRQ 125
           ||::   :   :: ::||:
/APLGSGARKPDEA---------PVESLDQLRDEPLQLAWEKPR------RQWPRR 266

VI---RAGLIYLYCQVHFDE---------------GKAVYLKLDL--LVNGV 162
 :| |||:   :   ::   : :::
LILLALGGLAAQYIAYHFDELARQDAYRPWFAQLCPEIGCTLPSKVDVDVQIRSSN 326

LEEFSA--------TAASSPGP--QLRLCQVSGLL----PLRPGSSL 200
 || |: |:| | |
HPEFSGALVVDAIIYNRASFSQPFPLLELRFADLNGHLIANRRFKPGEYL 381

ELIMINARY;   PRT;   213 AA.

rEMBLrel. 24, Created)
rEMBLrel. 24, Last sequence update)
rEMBLrel. 25, Last annotation update)
rotein.

vermitilis.
nobacteria; Actinobacteridae; Actinomycetales;
e; Streptomycetaceae; Streptomyces.
03;

N.A.
/ ATCC 31267 / NCIMB 12804 / NRRL 8165;
03; PubMed=11572948;
a H., Ishikawa J., Hanamoto A., Takahashi C.,
kahashi Y., Horikawa H., Nakazawa H., Osonoe T.,

ton K.A., Howland T.J., Wei M.-H., Ibegwam C.,
sh F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
dira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
B., McIntosh T.C., McLeod M.P., McPherson D.,
Ishina N.v., Mobarry C., Morris J., Moshrefi A.,
M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
lson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
ittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
ington K., Saunders R.D.C., Scheeler F., Shen H.,
n-Kiamos I., Simpson M., Skupski M.P., Smith T.,
ling A.C., Stapleton M., Strong R., Sun E.,
ctor C., Turner R., Venter E., Wang A.H., Wang X.,
sarman D.A., Weinstock G.M., Weissenbach J.,
Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
rg F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
rs E.W., Rubin G.M., Venter J.C.;
uence of Drosophila melanogaster.";
5-2195(2000).

.A.
Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
ayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.-
., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
ett V., Champe M., Davenport L.B., Dietz S.M.,
ise E., Doup L.E., Doyle C., Dresnek D., Farfan D.,
uck J., Galle R.F., Garg N.S., George R.A.,
ali M., Kruse D., Li P., Mattei A., Moshrefi A.,
Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
gas V., Park S., Patel S., Pfeiffer S.,
., Pittman G.S., Puri V., Richards S., Scheeler F.,
trong R., Sviriskas R., Tector C., Tyler D.,
Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
Drosophila melanogaster genome.";
2000) to the EMBL/GenBank/DDBJ databases.

.A.
y M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
ang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
gman C., Berman P., Carlson J.W., Celniker S.E.,
ale R., Emmert D., Frise E., de Grey A., Harris N.,
Marshall B., Millburn G., Richter J., Russo S.,
Smith E., Shu S., Smutniak F., Whitfield E.,
elbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
Drosophila melanogaster genome.";
2000) to the EMBL/GenBank/DDBJ databases.

.A.
niker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
2000) to the EMBL/GenBank/DDBJ databases.

.A.
2002) to the EMBL/GenBank/DDBJ databases.
CONTAINS 1 PDZ/DHR DOMAIN.
AAF55127.2; -.
PDR.
051304; CG31304.
C:cytoskeleton; IEA.
P:intracellular signaling cascade; IEA.
00299; Band 4.1.
01478; PDZ.
01202; WW_Rsp5_WWP.
PDZ; 1.
WW; 1.
B41; 1.
PDZ; 1.
WW; 1.
57; FERM_3; 1.
06; PDZ; 1.
59; WW_DOMAIN_1; 1.

```
DR   PROSITE; PS50020; WW_DOMAIN_2; 1.
SQ   SEQUENCE  1363 AA;  148691 MW;  876292CEC51FBE8D CRC64;  Length 1363;

Query Match             7.7%;   Score 89.5;  DB 5;
Best Local Similarity  21.7%;   Pred. No. 20;
Matches  49;  Conservative  27;  Mismatches  55;  Indels  95;  Ga

Qy    29  QEPSQELTAEDRREPPELNPQTESQDVVPFLEQLVRPRRSAPKGRKARPRR-AII
              ||:|| |                          |  : ::
Db  1040  EQPGQEEL------------QPPPRTPTTEQ------LSPPPARPPKSAELLQRYSPKKQVRI

Qy    87  YEVHPR--------------------------PGQDGAQAGVDGTVSGWEETKI-
             : ::                          ||:|| |
Db  1088  PVMQPQERRELCPPQLPPRGSPTLDGSQSSPTNAVSGPKKPPLPPIACRPRPSNGVN

Qy   116  SSS------PLRYDRQI-----------------------------GEFTVIRAGLYYLY(
             |||       :|| |                               ||| |: :||
Db  1148  SSSPGSAPPAHYSPPIPATVRLPHLNQANGTLPLLPKKPQQLHGEKLFIKNG------

Qy   144  HFDEGKAVYLKLDLLVNGVL------ALRCLEEFSATAASSPGQL 183
              ||||  | : :|| |            |: :::    : :|
Db  1200  HLIDGEALLAKTDVAMSGLLIKLDQVAAQC----SVAQAAGGGTSI 1241


RESULT 15
Q90WT9
ID   Q90WT9          PRELIMINARY;    PRT;   287 AA.
AC   Q90WT9;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   TNF-related apoptosis inducing ligand.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianin
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bridgham J.T., Johnson A.L.;
RT   "TNF-related apoptosis inducing ligand (TRAIL) expression in the h
RT   ovary.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY057941; AAL23702.1; -.
DR   HSSP; O35235; 1IQA.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005164; F:tumor necrosis factor receptor binding; IEA.
DR   GO; GO:0006955; P:immune response; IEA.
DR   InterPro; IPR006053; TNF_abc.
DR   InterPro; IPR006052; TNF_family.
DR   InterPro; IPR008983; TNF_like.
DR   InterPro; IPR003636; TNF_subf.
DR   Pfam; PF00229; TNF; 1.
DR   PRINTS; PR01234; TNECROSISFCT.
DR   ProDom; PD002012; TNF_subf; 1.
DR   SMART; SM00207; TNF; 1.
DR   PROSITE; PS50049; TNF_2; 1.
SQ   SEQUENCE  287 AA;  32092 MW;  DB06E1C95087B108 CRC64;

Query Match             7.7%;   Score 89;  DB 13;  Length 287;
Best Local Similarity  23.4%;   Pred. No. 2.9;
Matches  40;  Conservative  24;  Mismatches  65;  Indels  42;  Ga

Qy    84  AAHYEVHPR-PGQDGAQAGVDG------TVSGWEETKINSS-SPLRYDRQIGEFTV
              || ||:|    ||:|| |            |::| |: :||::  : |  ::
Db   127  SAHLIRPQNPAQDGSRRFGNLSQSCRHAITRWEDSTIHSHLQNITY--RDGRLRV

Qy   135  GLYYLYCQVHFD---------------EGKAVYLKLDLLVNGVLALRCLEEFS
              | :|    :                 || :| : :|| |  :  : :
Db   185  GKYYVVSQIYFRYSRDGAGARVSVPQLVQCINWKTSYSQPILLKGV---------

Qy   176  ASSPGPQ--LRLCQVSGLLPLRPGSSLRIRTLPWAHLKAAPPLTYFGLFQV 224
              ::  |    |:|| | ||| :| :::  : ||: :|:: || |:  :||
```

EAEYGLHALYQGGLFELKAGDELFVSVSSLAIDYSDAAASYFGAFRL 285

April 7, 2004, 17:46:43
secs

```
GenCore version 5.1.6
opyright (c) 1993 - 2004 Compugen Ltd.

n search, using sw model

il  7, 2004, 17:37:32 ; Search time 45.0884 Seconds
                         (without alignments)
                         1409.967 Million cell updates/sec

09-245-198A-2
2
LSLGLAIACLGLLLVVVSL.........PWAHLKAAPFLTYFGLFQVH 225

SUM62
op 10.0 , Gapext 0.5

6107 seqs, 282547505 residues

s satisfying chosen parameters:    1586107

th: 0
th: 2000000000

nimum Match 0%
ximum Match 100%
sting first 45 summaries

                    SUMMARIES
```

```
Geneseq_29Jan04:*
 geneseqp1980s:*
 geneseqp1990s:*
 geneseqp2000s:*
 geneseqp2001s:*
 geneseqp2002s:*
 geneseqp2003as:*
 geneseqp2003bs:*
 geneseqp2004s:*

the number of results predicted by chance to have a
than or equal to the score of the result being printed,
d by analysis of the total score distribution.
```

ry	ch	Length	DB	ID		Description
1.0		225	2	AAW47524	Aaw47524	Mus muscu
1.0		225	2	AAB07527	Aab07527	Amino aci
0.7		249	7	ADC97712	Adc97712	Murine FL
0.7		211	2	AAW93591	Aaw93591	Mouse TNR
0.8		249	2	AAW29745	Aaw29745	TNF relat
0.8		249	2	AAY09369	Aay09369	Human tum
0.8		249	3	AAY95338	Aay95338	Human PRO
0.8		249	4	AAB07526	Aab07526	Amino aci
0.8		249	4	AAE00891	Aae00891	Human TRE
0.8		249	5	AAU86129	Aau86129	Human PRO
0.8		249	6	ABR42315	Abr42315	Human TWE
0.8		249	7	ADC35206	Adc35206	Human TNF
0.8		284	2	AAW47525	Aaw47525	Homo sapi
0.8		208	4	AAW93590	Aaw93590	Human TNR
0.8		273	4	AAU03499	Aau03499	TWEAK ext
0.9		189	2	AAW29746	Aaw29746	TNF relat
0.9		189	4	AAE00892	Aae00892	Human UL4
0.0		146	4	AAE00895	Aae00895	Human TRE
0.6		406	5	AAU77717	Aau77717	Drosophil
0.5		409	5	AAU77718	Aau77718	Drosophil
0.3		325	2	ABB67553	Abb67553	Drosophil
0.2		294	2	AAW69956	Aaw69956	NF-xB rec
0.2		294	2	AAW68292	Aaw68292	NF-xB rec
0.2		294	2	AAE08737	Aae08737	Murine re
0.2		294	4	AAE04425	Aae04425	Murine re
26	106.5	9.2	294	4	AAE01992	Aae01992 Mur
27	106.5	9.2	294	5	AAB26102	Aae26102 Mou
28	106.5	9.2	294	7	ADB16986	Adb16986 Mur
29	106.5	9.2	294	7	ADC73000	Adc73000 Mur
30	106.5	9.2	294	7	ADC78266	Adc78266 Mur
31	103	8.9	220	4	AAB62340	Aab62340 Gpl
32	103	8.9	426	6	ABU39962	Abu39962 Prc
33	102.5	8.8	316	2	AAW83017	Aaw83017 Ost
34	102.5	8.8	316	2	AAW83194	Aaw83194 Hun
35	102.5	8.8	316	2	AAW59654	Aaw59654 Ami
36	102.5	8.8	316	3	AAY17874	Aay17874 Mur
37	102.5	8.8	316	3	AAY91024	Aay91024 Mou
38	102.5	8.8	316	3	AAY84418	Aay84418 Ami
39	102.5	8.8	316	5	AAY84419	Aay84419 Ami
40	102.5	8.8	316	5	AAU78289	Aau78289 Mou
41	102.5	8.8	316	6	ABR42071	Abr42071 Hun
42	102.5	8.8	316	6	ABB99477	Abb99477 Ami
43	102.5	8.8	316	6	ABU08463	Abu08463 Ami
44	102.5	8.8	316	6	ABR55560	Abr55560 Ami
45	99	8.5	234	4	AAB62339	Aab62339 Gp:

ALIGNMENTS

```
RESULT 1
AAW47524
ID   AAW47524 standard; protein; 225 AA.
XX
AC   AAW47524;
XX
DT   21-JUL-1998  (first entry)
XX
DE   Mus musculus tumour necrosis factor related ligand (TRELL).
XX
KW   TRELL; tumour necrosis factor related ligand; tnf; treatment; canc
KW   autoimmune disease; immune system; stimulation; suppression;
KW   graft rejection.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Domain          1..21
FT                   /note= "hydrophobic, transmembrane domain"
XX
PN   WO9805783-A1.
XX
PD   12-FEB-1998.
XX
PF   07-AUG-1997;    97WO-US013945.
XX
PR   07-AUG-1996;    96US-0023541P.
PR   18-OCT-1996;    96US-0028515P.
PR   18-MAR-1997;    97US-0040820P.
XX
PA   (BIOJ ) BIOGEN INC.
PA   (UVGE-) UNIV GENEVA FACULTY MEDICINE.
XX
PI   Chicheportiche Y, Browning JL;
XX
DR   WPI; 1998-145619/13.
DR   N-PSDB; AAV18599.
XX
PT   Tumour necrosis factor related ligand - useful for, e.g. treating
PT   auto-immune disease and immune responses to tissue grafts.
XX
PS   Claim 12; Page 48-50; 69pp; English.
XX
CC   The sequence is that of mouse tumour necrosis factor related liga
CC   (TRELL). TRELL or active fragments can be included with a carrier
CC   pharmaceutical compositions to treat cancer, autoimmune diseases
CC   immune responses to tissue grafts, or to stimulate or suppress th
CC   system. It is useful to screen for TRELL receptors, by labelling
```

bel and screening compositions for binding. Agents
ith TRELL-receptor binding can also be screened for, can
istered, optionally with interferon- gamma, to induce cell
t, suppress or alter immune responses (especially involving
rcinoma cells) involving a signal pathway between TRELL and
It's coding sequence can be used in gene therapy for TRELL-
ders in mammals (especially humans), e.g. tumours,
d inflammatory diseases or inherited genetic disorders, by
nto cells, and expressing, therapeutically effective amounts
e.g. a virus comprising a gene encoding TRELL. It may also
the preparation of prepare probes for screening
etic DNAs for TRELL-encoding sequences and for antisense

AA;

arity   100.0%; Score 1162; DB 2; Length 225;
           100.0%; Pred. No. 3.1e-112;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

GLALACLGLLLVVVSLGSWATLSAQEPSQEELTAEDRREPPELNPQTEESQDVVPF 60
GLALACLGLLLVVVSLGSWATLSAQEPSQEELTAEDRREPPELNPQTEESQDVVPF 60

VRPRRSAPKGRKARPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSSSPL 120
VRPRRSAPKGRKARPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSSSPL 120

QIGEFTVIRAGLYLYCQVHFDEGKAVYLKLDLLVNGVLALRCLEEFSATAASSPG 180
QIGEFTVIRAGLYLYCQVHFDEGKAVYLKLDLLVNGVLALRCLEEFSATAASSPG 180

LCQVSGLLPLRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 225
LCQVSGLLPLRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 225

dard; protein; 225 AA.

(first entry)

quence of a soluble recombinant murine TWEAK protein.

; immunological disorder; immune response; inflammation;
g agent; autoimmune disease; organ transplant rejection;
Host disease; GVHD; lymphoid cell malignancy; shock; tumour.

l.

2000WO-US001044.

99US-0116168P.

EN INC.

5036/41.

nd treating immune responses using modulators, especially
of TWEAK, TWEAK receptors and TWEAK ligands, useful for
. inflammation and graft versus host disease.

Fig 1; 45pp; English.

CC   The present sequence represents a TWEAK protein. The specificatio
CC   describes a method for preventing or treating an immunological di
CC   and/or inhibiting an immune response in an animal. The method may
CC   administering a TWEAK blocking agent. The method may be used for
CC   preventing and treating immune disorders associated with inappror
CC   expression and/or activity of TWEAK. These disorders include auto
CC   diseases, acute and chronic inflammation, organ transplant reject
CC   Graft-versus-Host disease (GVHD), lymphoid cell malignancies, sep
CC   other forms of shock, loss of immune responsiveness (as seen in h
CC   immunodeficiency virus (HIV) infections) and failure of the immur
CC   response to tumour growth
XX
SQ   Sequence 225 AA;

Query Match   100.0%; Score 1162; DB 3; Length 225;
Best Local Similarity   100.0%; Pred. No. 3.1e-112;
Matches 225; Conservative   0; Mismatches   0; Indels   0; G

Qy   1 VLSLGLALACLGLLLVVVSLGSWATLSAQEPSQEELTAEDRREPPELNPQTEESQI
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 VLSLGLALACLGLLLVVVSLGSWATLSAQEPSQEELTAEDRREPPELNPQTEESQI

Qy   61 LEQLVRPRRSAPKGRKARPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINS
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LEQLVRPRRSAPKGRKARPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINS

Qy 121 RYDRQIGEFTVIRAGLYLYCQVHFDEGKAVYLKLDLLVNGVLALRCLEEFSATAA
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 RYDRQIGEFTVIRAGLYLYCQVHFDEGKAVYLKLDLLVNGVLALRCLEEFSATAA

Qy 181 PQLRLCQVSGLLPLRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 225
        |||||||||||||||||||||||||||||||||||||||||||||
Db 181 PQLRLCQVSGLLPLRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 225

RESULT 3
ADC97712
ID   ADC97712 standard; protein; 249 AA.
XX
AC   ADC97712;
XX
DT   15-JAN-2004 (first entry)
XX
DE   Murine FL-TWEAK.
XX
KW   Murine; FL-TWEAK; TNF relatedness and weak ability to induce cel
KW   TNF; Tumour Necrosis Factor; TWEAK; fibrosis; cardiac disease;
KW   liver disease; lung disease; kidney disease; skin disease;
KW   skeletal muscle disease; adipose tissue disease;
KW   gastrointestinal tract disease; pancreatic disease;
KW   reproductive organ disease; neural disease; cartilage disease;
KW   bone disease; connective tissue disease; cellular death; hepatot
KW   dermatological; gastrointestinal; osteopathic.
XX
OS   Mus sp.
XX
PN   WO200308631l-A2.
XX
PD   23-OCT-2003.
XX
PF   09-APR-2003; 2003WO-US011350.
XX
PR   09-APR-2002; 2002US-0371611P.
XX
PA   (BIOJ) BIOGEN INC.
XX
PI   Burkly L, Jakubowski A, Zheng T, Hahm K;
XX
DR   WPI; 2003-845256/78.
DR   N-PSDB; ADC97713.
XX
PT   Treating a TWEAK-related condition, e.g. liver, gastrointestinal
PT   bone, pancreatic, cartilage or neural tissue condition in a subj

...nistering to the subject a TWEAK agonist or antagonist.

ID NO 1; 120pp; English.

...quence is murine transmembrane FL-TWEAK (TNF relatedness
...ty to induce cell death, where TNF is Tumour Necrosis
is a member of the TNF family. TWEAK agonists or
...e useful for treating a TWEAK-related condition, e.g.
iac disease; liver disease; lung disease; kidney disease;
skeletal muscle disease; adipose tissue disease;
...l disease; pancreatic disease; reproductive organ
...1 disease; cartilage disease; bone disease; connective
; cellular death; and a pathological condition of a tissue
WEAK receptor.

A;

...rity 100.0%;   Score 1162; DB 7;   Length 249;
...rity 100.0%;   Pred. No. 3.6e-112;
...nservative   0; Mismatches   0;   Indels   0;   Gaps   0;

LALACLGLLLVVVSLGSWATLSAQEPSQEELTAEDRREPPELNPQTEESQDVVPF   60
||||||||||||||||||||||||||||||||||||||||||||||||||||||
LALACLGLLLVVVSLGSWATLSAQEPSQEELTAEDRREPPELNPQTEESQDVVPF   84

RPRRSAPKGRKARPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSSSPL   120
||||||||||||||||||||||||||||||||||||||||||||||||||||||
RPRRSAPKGRKARPRRAIAAHYEVHPRPCGQDGAQAGVDGTVSGWEETKINSSSPL   144

IIGEFTVIRAGLYYLYCQVHFDEGKAVYLKLDLLVNGVLALRCLEEFSATAASPG   180
||||||||||||||||||||||||||||||||||||||||||||||||||||||
IGEFTVIRAGLYYLYCQVHFDEGKAVYLKLDLLVNGVLALRCLEEFSATAASPG   204

CQVSGLLPLRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH   225
|||||||||||||||||||||||||||||||||||||||
CQVSGLLPLRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH   249

...iard; protein; 211 AA.

...(first entry)

...otein.

...s factor receptor; signal transducer molecule; TNF; APO4;
...abnormality; gestational abnormallity; prostate cancer;
PO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
...main; immunogen; antibody preparation; breast carcinoma;
...se.

98WO-US018393.

97US-00924634.

WASHINGTON.

191/17.
425.

...rosis Factor family receptor polypeptides and ligands -
...agnosis and treatment of prostate cancer and developmental

---

PT   or gestational abnormalities.
XX
PS   Claim 40; Fig 13B; 156pp; English.
XX
CC   This invention describes isolated Tumor Necrosis Factor (TNF) fami...
CC   receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC   fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNF
CC   their active fragments. APO4 is useful for diagnosing prostate car
CC   determining levels of APO4 in an individual. Prostate cancer can a
CC   treated using APO4 selective binding agents linked to a therapeuti
CC   moiety. APO4 polypeptides are also useful for identifying selecti
CC   binding agents, useful in diagnosis/treatment of disease by bindi
CC   agents to the polypeptide/active fragment which is extracellular,
CC   expressed on the cell surface. The binding is preferably performe
CC   vivo. APO4 polypeptides/ active fragments are also useful for scr
CC   for agonists and antagonists by binding and observing the changer
CC   activity. Effective pharmacological agents useful in diagnosis or
CC   treatment of disease are also identified using APO4 polypeptides/
CC   fragments and APO4 signal transducer molecules that specifically
CC   with a cytoplasmic domain of APO4 and detecting a change in level
CC   activity. The method is performed in vivo or in vitro. APO polype
CC   are all useful as immunogens for preparing antibodies. APO4 is al
CC   useful for diagnosis/treatment of developmental or gestational
CC   abnormalities. APO8 was transfected to human breast carcinoma cel
CC   MCF-7, and induced apoptosis
XX
SQ   Sequence 211 AA;

Query Match   93.7%;   Score 1089; DB 2;   Length 211;
Best Local Similarity   99.5%;   Pred. No. 1.1e-104;
Matches 210; Conservative   0; Mismatches   1;   Indels   0; G

Qy   15 LVVVSLGSWATLSAQEPSQEELTAEDRREPPELNPQTEESQDVVPFLEQLVRPRRS
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 LVVVSLGSWATLSAQEPSQEELTAEDRREPPELNPQTEESQDVVPFLEQLVRPRRS

Qy   75 RKARPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSSSPLRYDRQIGEFT
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 RKARPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSSSPLRYDRQIGEFT

Qy 135 GLYYLYCQVHFDEGKAVYLKLDLLVNGVLALRCLEEFSATAASSPGPQLRLCQVSG
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 GLYYLYCQVHFDEGKAVYLKLDLLVNGVLALRCLEEFSATAASSPGPQLRLCQVSG

Qy 195 RPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 225
        ||||||||||||||||||||||||||||||
Db 181 RPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 211

RESULT 5
AAW29745
ID   AAW29745 standard; protein; 249 AA.
XX
AC   AAW29745;
XX
DT   27-OCT-1998 (first entry)
XX
DE   TNF related endothelium proliferative agent protein.
XX
KW   TNF; endothelium proliferative agent; TREPA; wound healing; cance
KW   tissue grafting; vascularisation; apoptosis; autoimmune; birth co
XX
OS   Homo sapiens.
XX
PN   WO9835061-A2.
XX
PD   13-AUG-1998.
XX
PF   12-FEB-1998;   98WO-US002859.
XX
PR   12-FEB-1997;   97US-00798692.
PR   10-FEB-1998;   98US-00021706.
XX

r LAB.

255/38.
613.

leic acid encoding TREPA - useful for diagnosis and
autoimmune disease, tumours and inflammation.

e 123-4; 142pp; English.

ed endothelium proliferative agent (TREPA), or its
agonists, are used to treat a deficit of TREPA, e.g. to
healing or tissue grafting, by promoting vascularisation,
e apoptosis for treating cancer and eliminating autoreactive
n adjunct to cancer chemotherapy or antiviral treatment.
s can also be used to target cytotoxic agents or for
ation of the corresponding receptor, the nucleic acid for
USED to transform tumour cells to render them more
TREPA and to screen for TREPA mimics. Ribozymes, antisense
ies or peptides, are used to treat TREPA-associated
, tumours and metastases (by inhibiting vascularisation),
or a wide range of autoimmune conditions, conditions
ormal stimulation of epithelial cells (e.g.
is), for birth control (inhibiting ovulation and placental
other angiogenic conditions (e.g. ulcers)

AA;

arity   87.8%;   Score 1020; DB 2;   Length 249;
88.8%;   Pred. No. 2.1e-97;
onservative   9;   Mismatches   16;   Indels   0;   Gaps   0;

LALACLGLLLVVVSLGSWATLSAQEPSQEELTAEDRREPPELNPQTEESQDVVPFL 61
|||||| ||||||| ||| |||| |||||||||||||||||| ||: :||
LALALACLGLLLLAVVSLGSRASLSAQEPAQEELVAEEDQDPSELNPQTESQDPAPFL 85

RPPRSAPKGRKARPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSSSPLR 121
||||||| |||| ||| |||||||||||||||||||||||||| :||||||||||
RPRRSAPKGRKTRARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLR 145

IGEFTVIRAGLYYLYCQVHFDEGKAVYLKLDLLVNGVLALRCLEEFSATAASSPGP 181
|||| |||||||||||||||||||||||||||| |||||||||||||||||| ||
IGEFIVTRAGLYYLYCQVHFDEGKAVYLKLDLLVDGVLALRCLEEFSATAASSLGP 205

CQVSGLLPLRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 225
||||| | ||||||||||||||||||||||||||||||||
CQVSGLLALRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 249

dard; protein; 249 AA.

(first entry)

necrosis factor Apo-3 ligand protein sequence.

; necrosis factor; Apo-3 ligand; lymphotoxin; apoptosis;
endent transcription; JNK/SAPK-dependent response; cancer.

98WO-US021407.
97US-0062037P.

PR   17-DEC-1997;   97US-0069862P.
XX
PA   (GETH) GENENTECH INC.
XX
PI   Ashkenazi AJ, Marsters SA, Pitti R;
XX
DR   WPI; 1999-287982/24.
DR   N-PSDB; AAX56000.
XX
PT   New human Apo3- ligand (a tumor necrosis factor) homologue.
XX
PS   Claim 1; Fig 1; 74pp; English.
XX
CC   The present sequence represents a human tumour necrosis factor (T
CC   lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has
CC   cytostatic activity. Apo-3 ligand can be used to induce apoptosis
CC   mammalian cancer cells, to induce NF-kappaB-dependent transcripti
CC   to induce JNK/SAPK-dependent responses in mammalian cells
XX
SQ   Sequence 249 AA;

Query Match   87.8%;   Score 1020; DB 2;   Length 249;
Best Local Similarity   88.8%;   Pred. No. 2.1e-97;
Matches 199; Conservative   9; Mismatches   16;   Indels   0;   C

Qy   2 LSLGLALACLGLLLVVVSLGSWATLSAQEPSQEELTAEDRREPPELNPQTEESQDV
          | |||||||||||| ||||||||||| |||||| |||||||||||||||||||||
Db   26 LGLGLALACLGLLLLAVVSLGSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQDF

Qy   62 EQLVRPRRSAPKGRKARPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSS
          :|| |||||| |||| ||| ||||||||||||||||||||||||||||| :||||
Db   86 NRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSS

Qy   122 YDRQIGEFTVIRAGLYYLYCQVHFDEGKAVYLKLDLLVNGVLALRCLEEFSATAAS
          |: |||| |||||||||||||||||||||||||||||||| |||||||||||||||
Db   146 YNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYLKLDLLVDGVLALRCLEEFSATAAS

Qy   182 QLRLCQVSGLLPLRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 225
          |||||||||| | |||||||||||||||||||||||||||||||
Db   206 QLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 249

RESULT 7
AAY95338
ID   AAY95338 standard; protein; 249 AA.
XX
AC   AAY95338;
XX
DT   25-SEP-2000 (first entry)
XX
DE   Human PRO207 antitumour protein.
XX
KW   PRO207; human; antitumour; tumour; therapy; cytostatic; breast ca
KW   ovarian cancer; renal cancer; colorectal cancer; uterine cancer;
KW   prostate cancer; lung cancer; bladder cancer;
KW   central nervous system cancer; melanoma; leukaemia; neoplasm.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   Peptide   1. .40
FT   /label= Signal_peptide
FT   Modified-site 10. .14
FT   /note= "amidation"
FT   Peptide   24. .35
FT   /note= "prokaryotic membrane lipoprotein lipid"
FT   Modified-site 27. .33
FT   /note= "N-myristoylation"
FT   Modified-site 29. .35
FT   /note= "N-myristoylation"
FT   Modified-site 36. .42
FT   /note= "N-myristoylation"
FT   Protein   41. .249

```
Qy    122  YDRQIGEFTVIRAGLYYLYCQVHFDEGKAVYLKLDLLVNGVLALRCLEEFSATAAS
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    146  YNRQIGEFIVTRAGLYYLYLYCQVHFDEGKAVYLKLDLLVDGVLAIRCLEEFSATAAS

Qy    182  QLRLCQVSGLLPLRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH  225
              ||||||||||||||||||||||||||||||||||||||||||||||
Db    206  QLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH  249
```

RESULT 8
AAB07526
ID   AAB07526 standard; protein; 249 AA.
XX
AC   AAB07526;
XX
DT   20-OCT-2000 (first entry)
XX
DE   Amino acid sequence of a soluble recombinant human TWEAK protein.
XX
KW   TWEAK protein; immunological disorder; immune response; inflammati;
KW   TWEAK blocking agent; autoimmune disease; organ transplant reject;
KW   Graft-versus-Host disease; GVHD; lymphoid cell malignancy; shock;
XX
OS   Homo sapiens.
XX
PN   WO200042073-A1.
XX
PD   20-JUL-2000.
XX
PF   14-JAN-2000; 2000WO-US001044.
XX
PR   15-JAN-1999; 99US-0116168P.
XX
PA   (BIOJ) BIOGEN INC.
XX
PI   Rennert P;
XX
DR   WPI; 2000-476036/41.
XX
PT   Preventing and treating immune responses using modulators, especi;
PT   antibodies, of TWEAK, TWEAK receptors and TWEAK ligands, useful f;
PT   treating e.g. inflammation and graft versus host disease.
XX
PS   Disclosure; Fig 1; 45pp; English.
XX
CC   The present sequence represents a TWEAK protein. The specification
CC   describes a method for preventing or treating an immunological di
CC   and/or inhibiting an immune response in an animal. The method comp
CC   administering a TWEAK blocking agent. The method may be used for
CC   preventing and treating immune disorders associated with inapprop
CC   expression and/or activity of TWEAK. These disorders include auto
CC   diseases, acute and chronic inflammation, organ transplant reject
CC   Graft-versus-Host disease (GVHD), lymphoid cell malignancies, sep
CC   other forms of shock, loss of immune responsiveness (as seen in h
CC   immunodeficiency virus (HIV) infections) and failure of the immun
CC   response to tumour growth
XX
SQ   Sequence 249 AA;

Query Match   87.8%; Score 1020; DB 3; Length 249;
Best Local Similarity   88.8%; Pred. No. 2.1e-97;
Matches 199; Conservative   9; Mismatches 16; Indels   0; G

```

```
Qy   2 LSLGLALACLGLLLVVVSLGSWATLSAQEPSQEELTAEDRREPPELNPQTEESQDV
              ||||||||||||||||||||||||||| |: :: : |||||||||||||||||||||
Db   26 LGLGLALACLGLLLAVVSLGSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQDP.

Qy   62 EQLVRPRRSAPKGRKARPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSS
              ||| | || || :: ||||||||||||||||||||||||||||||| ||||
Db   86 NRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSS

Qy   122 YDRQIGEFTVIRAGLYYLYCQVHFDEGKAVYLKLDLLVNGVLALRCLEEFSATAAS
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        146  YNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYLKLDLLVDGVLALRCLEEFSATAAS
               |:||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy        182  QLRLCQVSGLLPRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH  225
               ||||||||||||||||||||||||||||||||||||||||||||

Db        206  QLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH  249
```

```
RESULT 10
AAU86129
ID   AAU86129 standard; protein; 249 AA.
XX
AC   AAU86129;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human PRO207 polypeptide.
XX
KW   Human; PRO; benign tumour; malignant tumour; lymphoid malignancy;
KW   leukaemia; neuronal disorder; stromal disorder; blastocoelic disc
KW   inflammatory disorder; immune disorder; angiogenic disorder; cyto
KW   neuroprotective.
XX
OS   Homo sapiens.
XX
PN   WO200153486-A1.
XX
PD   26-JUL-2001.
XX
PF   11-FEB-2000; 2000WO-US003565.
XX
PR   08-MAR-1999;      99WO-US005028.
PR   11-MAR-1999;      99US-0123972P.
PR   11-MAY-1999;      99US-0133459P.
PR   02-JUN-1999;      99WO-US012252.
PR   22-JUN-1999;      99US-0140650P.
PR   22-JUN-1999;      99US-0140653P.
PR   20-JUL-1999;      99US-0144758P.
PR   26-JUL-1999;      99US-0145698P.
PR   28-JUL-1999;      99US-0146222P.
PR   17-AUG-1999;      99US-0149395P.
PR   31-AUG-1999;      99US-051688P.
PR   01-SEP-1999;      99WO-US020111.
PR   15-SEP-1999;      99WO-US021090.
PR   30-NOV-1999;      99WO-US028313.
PR   01-DEC-1999;      99WO-US028301.
PR   05-JAN-2000;      99WO-US028634.
PR   05-JAN-2000;      2000WO-US000219.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Ashkenazi AJ,   Goddard A,   Godowski PJ,   Gurney AL,   Hillan KJ;
PI   Marsters SA,    Pan J,   Pitti RM,   Roy MA,   Smith V,   Stone DM;
PI   Watanabe CK,    Wood WI;
XX
DR   WPI; 2002-205567/26.
DR   N-PSDB; ABK40255.
XX
PT   Thirty five nucleic acids encoding PRO polypeptides, useful for t
PT   benign or malignant tumors, leukemias and lymphoid malignancies,
PT   inflammatory, angiogenic and immunologic disorders.
XX
PS   Claim 61; Fig 4; 302pp; English.
XX
CC   The present invention relates to the isolation of novel human PRO
CC   polypeptides and the polynucleotide sequences encoding them. The
CC   polypeptides, agonists, antagonists or anti-PRO antibodies are us
CC   treating benign or malignant tumours (e.g. renal, kidney, bladder
CC   breast, etc), leukaemias and lymphoid malignancies, other disorde
CC   as neuronal, glial, astrocytal, hypothalamic, glandular, macropha
CC   stromal and blastocoelic disorders, inflammatory, immune and angi
CC   disorders. The polynucleotide sequences are also useful in gene t
CC   AAU86128-AAU86162 represent the human PRO polypeptides of the inv
```

```
Db        IGEFIVTRAGLYYLYCQVHFDEGKAVYLKLDLLVDGVLALRCLEEFSATAASLGP  205

Qy        CQVSGLLPRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH  225
          |||||| |||||||||||||||||||||||||||||||||
Db        CQVSGLLALRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH  249
```

```
dard; protein; 249 AA.


(first entry)

TNF related endothelium proliferative agent).

necrosis factor; TNF; angiogenesis; wound healing; TREPA;
ndothelium proliferative agent; tumour; metastasis;
nerary.


Location/Qualifiers
98  .249
     /label= Extracellular_domain


760/29.
350.


98US-0105343.


97US-00798692.
98US-00021706.


T LAB.


75-76; 53pp; English.

nvention relates to extracellular signal molecules,
members of tumour necrosis factor (TNF) family molecules
TREPA (TNF related endothelium proliferative agent).
gically active TREPA are used to treat TREPA-associated
ours or metastases. TREPA is used for inducing angiogenesis
promoting wound healing and for vascularising grafted tissue
l grafting and to promote tissue grafts. The present amino
is clone ID #690050 human TREPA
```

ogenesis in mammal at desired sites for promoting wound
ministering soluble fragment of extracellular domain of
s factor related endothelium proliferative agent protein.

```
AA;

       87.8%;  Score 1020;  DB 4;  Length 249;
arity  88.8%;  Pred. No. 2.1e-97;
onservative   9;  Mismatches   16;  Indels   0;  Gaps   0;
```

```
:LALACLGLLLVVSLGSWATLSAQEPSQEELTAEDRREPPELNPQTEESQDVVPFL  61
 ||||||||||||||||||||||||||||||||||||||:::||||||||| |||
:LALACLGLLLAVVSLGSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQDPAPFL  85

RPRRSAPKGRKARPRRATAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSSSPLR  121
|||||||||||:||||||||||||||||||||||||||| :||||||||||||||
RPRRSAPKGRKTRARRATAAHYEVHPRPGQDGAQAGVDGTVSGWEEAARINSSSPLR  145

)IGEFTVIRAGLYYLYCQVHFDEGKAVYLKLDLLVNGVLALRCLEEFSATAASSPGP  181
```

A;
  87.8%; Score 1020; DB 5; Length 249;
rity 88.8%; Pred.No.2.1e-97;
nservative 9; Mismatches 16; Indels 0; Gaps 0;

```
ALACLGLLLVVVSLGSWATLSAQEPSQEELTAEDRREPPELNPQTEESQDVVPFL 61
|||||||||||||||||| |:|||||||||||:|:|||| : |||||||||
ALACLGLLLAVVSLGSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQDPAPFL 85

PRRSAPKGRKARPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSSSPLR 121
|||||||||||||||||||||||||||||||||||||||  :|||
PRRSAPKGRKTRARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLR 145

GEFTVIRAGLYYLYCQVHFDEGKAVYLKDLLVNGVLALRCLEEFSATAASSPGP 181
||||:|||||||||||||||||||||||||||:|||||||||||||
GEFIVTRAGLYYLYCQVHFDEGKAVYLKDLLVDGVLALRCLEEFSATAASSLGP 205

QVSGLLLPLRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 225
||||||||||||||||||||||||||||||||||
QVSGLLALRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 249
```

lard; protein; 249 AA.

first entry)

otein.

tumour necrosis factor; ligand; cytostatic;
r; osteopathic.

2.

002WO-US023782.
001US-0307838P.

GENOME SCI INC.

osen CA,

59/40.
01.

imeric complex having a first polypeptide member of the
factor (TNF) ligand family, and a second different member
family, useful for treating cancer, osteoporosis or an
ease.

ge 368-369; 388pp; English.

quence is the protein sequence for human TWEAK protein. The
tes to compositions comprising heterotrimeric complexes of
s factor (TNF) ligand family members, and their use in the
vention and treatment of disease. In one embodiment, the
complex comprises full-length or extracellular portions of
-length or extracellular portions of other TNF ligand
, preferably VEGI or VEGI-SV. The heterotrimeric complexes
on are useful for treating an autoimmune disease, cancer or
and particularly for inhibiting cancer cell proliferation,
ell proliferation, or inducing apoptosis of T cells

A;

Query Match   87.8%; Score 1020; DB 6; Length 249;
Best Local Similarity 88.8%; Pred. No. 2.1e-97;
Matches 199; Conservative 9; Mismatches 16; Indels 0; Ga

```
Qy    2 LSIGLALACLGLLLVVVSLGSWATLSAQEPSQEELTAEDRREPPELNPQTEESQDV
          | |||||||||||||||| |:|||||||||||::: |||| : :|||||||||||
Db   26 LGLGLALACLGLLLAVVSLGSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQDPA

Qy   62 EQLVRPRRSAPKGRKARPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSS
          :||||||||||||||| |||||||||||||||||||||||||||||| :|||
Db   86 NRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSS

Qy  122 YDRQIGEFTVIRAGLYYLYCQVHFDEGKAVYLKDLLVNGVLALRCLEEFSATAAS
          |:||||||:|:|||||||||||||||||||||||||:|||||||||||||||||
Db  146 YNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYLKDLLVDGVLALRCLEEFSATAAS

Qy  182 QLRLCQVSGLLLPLRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 225
          |||||||||||||| |||||||||||||||||||||||||||||
Db  206 QLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 249
```

RESULT 12
ADC35206
ID   ADC35206 standard; protein; 249 AA.
XX
AC   ADC35206;
XX
DT   18-DEC-2003 (first entry)
XX
DE   Human TNF ligand family member #12.
XX
KW   human; tumour necrosis factor; TNF ligand; endokine alpha;
KW   excessive bone resorption disorder; osteoporosis; Paget's disease.
KW   arterial calcification.
XX
OS   Homo sapiens.
XX
PN   US2003100074-A1.
XX
PD   29-MAY-2003.
XX
PF   15-AUG-2002; 2002US-00218547.
XX
PR   16-AUG-2001; 2001US-0312542P.
PR   30-OCT-2001; 2001US-0330761P.
XX
PA   (YUGG/) YU G.
PA   (NIJJ/) NI J.
PA   (ROSE/) ROSEN C A.
PA   (NARD/) NARDELLI B.
XX
PI   Yu G, Ni J, Rosen CA, Nardelli B;
XX
DR   WPI; 2003-696072/66.
DR   N-PSDB; ADC35205.
XX
PT   New Endokine alpha gene useful for preparing a composition for tre
PT   disease associated with excessive or insufficient bone resorption
PT   osteoporosis, Paget's disease or arterial calcification.
XX
PS   Disclosure; SEQ ID NO 24; 145pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule encodi
CC   tumour necrosis factor family ligand. A composition comprising the
CC   isolated antibody or its fragment is used for treating an individu
CC   need of decreased level of endokine alpha activity. The endokine a
CC   polypeptide present in a heterotrimeric complex is used for treat
CC   individual having a disorder associated with excessive bone resor
CC   e.g. osteoporosis, Paget's disease or arterial calcification. Tre
CC   individual having a disorder associated with insufficient bone res
CC   comprises administering an endokine alpha antagonist, which is the
CC   antibody that binds specifically to endokine alpha polypeptide. Th

CC   death or treat, suppress or alter immune responses (especially i
CC   human adenocarcinoma cells) involving a signal pathway between T
CC   its receptor. It's coding sequence can be used in gene therapy f
CC   related disorders in mammals (especially humans), e.g. tumours,
CC   autoimmune and inflammatory diseases or inherited genetic disord
CC   introducing into cells, and expressing, therapeutically effectiv
CC   of a vector, e.g. a virus comprising a gene encoding TRELL. It m
CC   be of use in the preparation of prepare probes for screening
CC   natural/synthetic DNAs for TRELL-encoding sequences and for anti
CC   therapy
XX
SQ   Sequence 284 AA;

Query Match   87.8%;   Score 1020;   DB 2;   Length 284;
Best Local Similarity   88.8%;   Pred. No. 2.6e-97;
Matches   199; Conservative   9; Mismatches   16;   Indels   0;   (

Qy   2 LSLGLALACLGLLLVVSLGSWATLSAQEPSQEELTAEDRREPPELNPQTEESQD
              |||||||||||||||||||||||||||:||||||||| :|||||||||||||||
Db   61 LGLGLALACLGLLLAVVSLGSRASLSAQEPAQELVAEEDQDPSELNPQTEESQD

Qy   62 EQLVRPRRSAPKGRKARPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEWTKINS
              |:|||||||||||||||||||||||||||||||||||||||||||| :|||
Db   121 NRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINS

Qy   122 YDRQIGEFTVIRAGLYYLYCQVHFDEGKAVYLKDLLVNGVLALRCLEEFSATAA
              |:||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 YNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYLKDLLVDGVLALRCLEEFSATAA

Qy   182 QLRLCQVSGLLPLRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 225
              |||||||||||||||||||||||||||||||||||||||||||
Db   241 QLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 284

RESULT 14
AAW93590
ID   AAW93590 standard; protein; 208 AA.
XX
AC   AAW93590;
XX
DT   18-JUN-1999 (first entry)
XX
DE   Human TNRL3 protein.
XX
KW   Tumour necrosis factor receptor; signal transducer molecule; TNF
KW   developmental abnormality; gestational abnormality; prostate c
KW   APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy;
KW   cytoplasmic domain; immunogen; antibody preparation; breast carc
KW   apoptosis; human.
XX
OS   Homo sapiens.
XX
PN   WO9911791-A2.
XX
PD   11-MAR-1999.
XX
PF   04-SEP-1998;   98WO-US018393.
XX
PR   05-SEP-1997;   97US-00924634.
XX
PA   (UNIW) UNIV WASHINGTON.
XX
PI   Chaudhary PM;
XX
DR   WPI; 1999-205191/17.
DR   N-PSDB; AAX23424.
XX
PT   New Tumor Necrosis Factor family receptor polypeptides and ligan
PT   useful for diagnois and treatment of prostate cancer and develo
PT   or gestational abnormalities.
XX
PS   Claim 40; Fig 13A; 156pp; English.
XX

nce represents the amino acid sequence of a tumour necrosis
ligand.
AA;

     87.8%;   Score 1020; DB 7;   Length 249;
arity 88.8%;   Pred. No. 2.1e-97;
onservative   9; Mismatches   16;   Indels   0;   Gaps   0;

LALACLGLLLVVSLGSWATLSAQEPSQEELTAEDRREPPELNPQTEESQDVVPFL 61
||||||||||||||| ||||||||||| :||||||||| |:|| ||||||||| ||
LALACLGLLLAVVSLGSRASLSAQEPAQELVAEEDQDPSELNPQTEESQDPAPFL 85

RPRRSAPKGRKARPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSSSPLR 121
|||||||||||| ||||||||||||||||||||||||||||||| :|||||||
RPRRSAPKGRKTRARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLR 145

DIGEFTVIRAGLYYLYCQVHFDEGKAVYLKDLLVNGVLALRCLEEFSATAASSPGP 181
||||| ||||||||||||||||||||||||||||| |||||||||||||||||||
DIGEFIVTRAGLYYLYCQVHFDEGKAVYLKDLLVDGVLALRCLEEFSATAASSLGP 205

CQVSGLLPLRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 225
||||||| ||||||||||||||||||||||||||||||||
CQVSGLLALRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 249

dard; protein; 284 AA.

(first entry)

tumour necrosis factor related ligand (TRELL).

: necrosis factor related ligand; tnf; treatment; cancer;
sease; immune system; stimulation; suppression;
on.

97WO-US013945.

96US-0023541P.
96US-0028515P.
97US-0040820P.

N INC.
GENEVA FACULTY MEDICINE.

e Y, Browning JL;

619/13.
600.

sis factor related ligand - useful for, e.g. treating cancer,
disease and immune responses to tissue grafts.

ge 50-51; 69pp; English.

is that of human tumour necrosis factor related ligand
L or active fragments can be included with a carrier in
al compositions to treat cancer, autoimmune diseases or
ases to tissue grafts, or to stimulate or suppress the immune
s useful to screen for TRELL receptors, by labelling with a
abel and screening compositions for binding. Agents
with TRELL-receptor binding can also be screened for, can
istered, optionally with interferon-gamma, to induce cell

XX
PF   19-DEC-2000; 2000WO-US034755.
XX
PR   20-DEC-1999; 99US-0172878P.
PR   10-MAY-2000; 2000US-0203347P.
XX
PA   (IMMV) IMMUNEX CORP.
XX
PI   Wiley SR;
XX
DR   WPI; 2001-417975/44.
DR   N-PSDB; AAS03964.
XX
PT   Modulating angiogenesis in a mammal for treating diseases mediated
PT   angiogenesis, e.g. solid tumors and vascular deficiencies of cardi
PT   peripheral tissue, by administering antagonist or agonist of TWEAK
PT   receptor.
XX
PS   Example 1; Page 41; 46pp; English.
XX
CC   The sequence represents a a fusion protein encoded by the expressi
CC   vector pDC409-LZ-TWEAK. The fusion protein comprises a growth horm
CC   leader, a leucine zipper multimerisation domain, and the extracel
CC   domain of human TWEAK. The fusion protein was used in the isolatio
CC   human TWEAK receptor (TWEAKR) -expressing clones from a COS cell hu
CC   cDNA library. The TWEAK protein is a member of the tumour necrosis
CC   (TNF) family and induces angiogenesis. TWEAKR may therefore be use
CC   screen for and develop TWEAKR agonists and antagonists for the mod
CC   of angiogenesis, to be used in the treatment and diagnosis of huma
CC   disease. The disorders mediated by angiogenesis include ocular dis
CC   characterised by ocular neovascularisation such as diabetic retino
CC   neovascular glaucoma, retinoblastoma, retinopathy of prematurity,
CC   retrolental fibroplasia, rubeosis, uveitis, macular degeneration a
CC   corneal graft neovascularisation, and inflammatory diseases such a
CC   arthritis, rheumatism and psoriasis. Other treatable diseases incl
CC   malignant and metastatic conditions such as sarcomas and carcinoma
CC   benign tumours and preneoplastic conditions, myocardial angiogenes
CC   haemophilic joints, scleroderma, vascular adhesions, atherosclerot
CC   plaque neovascularisation, telangiectasia, wound granulation, coro
CC   atherosclerosis, peripheral atherosclerosis and ischaemia
XX
SQ   Sequence 273 AA;

Query Match   81.8%; Score 951; DB 4; Length 273;
Best Local Similarity 88.9%; Pred. No. 3.6e-90;
Matches   184; Conservative   9; Mismatches   14; Indels   0; Ga

Qy   19 SLGSWATLSAQEPSQEELTAEDRREPPELNPQTEESQDVVPFLEQLVRPRRSAPKGI
             ||||||:||||||||||||||:.::||||||||||||||||| |||:|||||||||||
Db   67 SLGSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGI

Qy   79 PRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSSSPLRYDRQIGEFTVIRAC
             |||:||||||||||||||||||||||||||| :|||||||||||||:|||||||||
Db 127 ARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAC

Qy 139 LYCQVHFDEGKAVYLKLDLLVNGVLALRCLEEFSATAASSPGPQLRLCQVSGLLPLF
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 187 LYCQVHFDEGKAVYLKLDLLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALF

Qy 199 SLRIRTLPWAHLKAAPFLTYFGLFQVH 225
             |||||||||||||||||||||||||
Db 247 SLRIRTLPWAHLKAAPFLTYFGLFQVH 273

Search completed: April 7, 2004, 17:44:45
Job time : 48.0884 secs

---

describes isolated Tumor Necrosis Factor (TNF) family
eptides: APO4, APO6, APO8 and APO9 or their active
isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
ragments. APO4 is useful for diagnosing prostate cancer by
vels of APO4 in an individual. Prostate cancer can also be
APO4 selective binding agents linked to a therapeutic
olypeptides are also useful for identifying selective
, useful in diagnosis/treatment of disease by binding of
polypeptide/active fragment which is extracellular, or
he cell surface. The binding is preferably performed in
ypeptides/ active fragments are also useful for screening
nd antagonists by binding and observing the changer in APO4
ctive pharmacological agents useful in diagnosis or
isease are also identified using APO4 polypeptides/active
APO4 signal transducer molecules that specifically interact
smic domain of APO4 and detecting a change in level of APO4
method is performed in vivo or in vitro. APO polypeptides
as immunogens for preparing antibodies. APO4 is also
gnosis/treatment of developmental or gestational
APO8 was transfected to human breast carcinoma cell line
uced apoptosis

A;

          81.8%; Score 951; DB 2; Length 208;
rity   88.9%; Pred. No. 2.5e-90;
nservative   8; Mismatches   15; Indels   0; Gaps   0;

WATLSAQEPSQEELTAEDRREPPELNPQTEESQDVVPFLEQLVRPRRSAPKGRKA 77
|||||:|| ::| :.::|||||||||||||||||||||| |||:|||||||||||
RASLSAQEPAQEELVAEEDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKT 60

IAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSSSPLRYDRQIGEFTVIRAGLY 137
|||||||||||||||||||||||||| :|||||||||||||:|||||||||||
IAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYRYRQIGEFIVTRAGLY 120

VHFDEGKAVYLKLDLLVNGVLALRCLEEFSATAASSPGPQLRLCQVSGLLPLRPG 197
|||||||||||||||||||||||||||||||||||||||||||| ||||
VHFDEGKAVYLKLDLLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPG 180

RTLPWAHLKAAPFLTYFGLFQVH 225
||||||||||||||||||||||
RTLPWAHLKAAPFLTYFGLFQVH 208

ard; protein; 273 AA.

first entry)

lular domain-containing fusion protein.

lular domain; tumour necrosis factor; TNF; angiogenesis;
ularisation; diabetic retinopathy; neovascular glaucoma;
; retinopathy of prematurity; retrolental fibroplasia;
tis; macular degeneration; arthritis; rheumatism;
neovascularisation; psoriasis; metastatic condition;
ur; sarcoma; carcinoma; benign tumour; haemophilic joint;
condition; myocardial angiogenesis; wound granulation;
ascular adhesion; telangiectasia; ischaemia; human;
c plaque neovascularisation; coronary atherosclerosis;
erosclerosis; pDC409-LZ-TWEAK; TWEAK receptor; TWEAKR;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

ic search, using sw model

ril 7, 2004, 17:30:19 ; Search time 3328.48 Seconds
                         (without alignments)
                         12318.149 Million cell updates/sec

-09-245-198A-3
73
atgtcattgttagactttga........gacaaaatgttgataaatgg 1373

ENTITY_NUC
pop 10.0 , Gapext 1.0

513289 seqs, 14931090276 residues

:s satisfying chosen parameters: 55026578

gth: 0
gth: 2000000000

inimum Match 0%
aximum Match 100%
sting first 45 summaries

ST:*
em_estba:*
em_esthum:*
em_estin:*
em_estmu:*
em_estov:*
em_estpl:*
em_estro:*
em_htc:*
gb_est1:*
gb_est2:*
gb_htc:*
gb_est3:*
gb_est4:*
gb_est5:*
em_estfun:*
em_estcom:*
em_gss_hum:*
em_gss_inv:*
em_gss_pln:*
em_gss_vrt:*
em_gss_mam:*
em_gss_mus:*
em_gss_pro:*
em_gss_rod:*
em_gss_phg:*
em_gss_vrl:*
gb_gss1:*
gb_gss2:*

the number of results predicted by chance to have a
than or equal to the score of the result being printed,
d by analysis of the total score distribution.

SUMMARIES

ry
ch Length DB ID   Description
------ ------ -- -------- -----------
5.2   776   13 BX090012 BX090012 BX090012
5.8   834   12 BI766766 BI766766 603056866
4.9   948   13 BQ707185 BQ707185 AGENCOURT
4.4   963   13 BQ671259 BQ671259 AGENCOURT

						Accession
5	743.4	54.1	777	12	BI819200	BI819200
6	725.4	52.8	1071	12	BM921213	BM921213
7	701.4	51.1	731	12	BI871711	BI871711
8	688	50.1	828	12	BI596681	BI596681
9	677.4	49.3	728	12	BI870393	BI870393
10	660.4	48.1	697	14	CF126932	CF126932
11	656.6	47.8	666	14	CF126539	CF126539
12	584.4	42.6	609	14	BI966060	BI966060
12	581.4	42.3	1027	28	AF163779	AF163779
14	567	41.3	567	14	CA396679	CA396679
15	562	40.9	940	13	BQ884231	BQ884231
16	560.6	40.8	2237	11	AK044387	AK044387
17	545	39.7	545	14	CB141389	CB141389
18	544.4	39.7	575	14	CA413067	CA413067
19	538.8	39.2	569	14	CB529199	CB529199
20	535.6	39.0	569	13	BU631264	BU631264
21	534.6	38.9	568	12	BM971606	BM971606
22	531.6	38.7	910	12	BG110063	BG110063
23	516.8	37.6	531	12	BI824443	BI824443
24	494.2	36.0	824	14	CB998034	CB998034
25	493.6	36.0	918	13	BF577781	BF577781
26	462	33.6	474	13	BU951915	BU951915
27	458.2	33.4	1033	11	AK020909	AK020909
28	454.4	33.1	561	10	AW763237	AW763237
29	444	32.3	951	13	BQ674188	BQ674188
30	442.8	32.3	587	12	BG686319	BG686319
31	440.2	32.1	456	12	BI966255	BI966255
32	436.8	31.8	440	12	BM128059	BM128059
33	431.6	31.4	538	10	BF821434	BF821434
34	404	29.4	436	28	AQ890280	AQ890280
35	399.4	29.1	785	12	BI762908	BI762908
36	399.2	29.1	665	13	BY742288	BY742288
37	399	29.1	892	14	CB204861	CB204861
38	394.6	28.7	543	29	CG565104	CG565104
39	394	28.7	413	9	AI422796	AI422796
40	389.2	28.3	698	12	BI906850	BI906850
41	381	27.7	1064	14	CF994566	CF994566
42	379.6	27.6	692	13	BY748962	BY748962
43	374.2	27.3	416	9	AI291866	AI291866
44	372.8	27.2	1319	14	CF594233	CF594233
45	372.6	27.1	584	10	AW917574	AW917574

ALIGNMENTS

RESULT 1
BX090012
LOCUS   BX090012   776 bp   mRNA   linear   EST 23
DEFINITION BX090012 Soares breast 2NbHBst Homo sapiens cDNA clone
            IMAGp998E15243 ; IMAGE:154742, mRNA sequence.
ACCESSION   BX090012
VERSION   BX090012.1 GI:27821952
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 776)
REFERENCE   Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Pete
AUTHORS   Radelof,U., Schneider,D. and Korn,B.
TITLE   Human UnigeneSet - RZPD3
JOURNAL   Unpublished (2003)
COMMENT   Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
            RZPD; IMAGp998E15243.
            RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
            Human UnigeneSet - RZPD3 (RZPDLIB No.972)
            http://www.rzpd.de/CloneCards/cgi-
            bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany

49 30 32639 101
49 30 32639 111
pd.de
lone is available royalty-free from R2PD;
t RZPD (clone@rzpd.de) for further information. Seg primer:
Primer sequence: TTTCACACAGAAACAGCTATGAC.
  Location/Qualifiers
  1..776
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGp998E15243 ; IMAGE:154742"
      /sex="Female"
      /dev_stage="adult"
      /lab_host="DH10B (ampicillin resistant)"
      /clone_lib="Soares breast 2NbHBst"
      /note="Organ: breast; Vector: pT7T3D (Pharmacia) with a
      modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
      strand cDNA was primed with a Not I - oligo(dT) primer [5'
      TGTTACCAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTT 3'],
      double-stranded cDNA was ligated to Eco RI adaptors
      (Pharmacia), digested with Not I and cloned into the Not I
      and Eco RI sites of a modified pT7T3 vector (Pharmacia).
      Library went through one round of normalization to a Cot =
      230. Library constructed by Bento Soares and M.Fatima
      Bonaldo."

arity   56.2%;   Score 772; DB 13;   Length 776;
99.7%;   Pred. No. 4.9e-135;
onservative   0; Mismatches   2; Indels   0; Gaps   0;

GTGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTCGTGGATGG 659
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTCGTGGATGG 60

CTGGCCCTGCGCTGCCTGGAGGAATTCTCAGCCACTGCGGGCCAGTTCCCTCGGGCC 719
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTGGCCCTGCGCTGCCTGGAGGAATTCTCAGCCACTGCGGGCCAGTTCCCTCGGGCC 120

CTCCGCCTCTGCCCAGGTGTCTGGGCTGTTGGCCCTGCGCGCCAGGTCCTCCCTGCG 779
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTCCGCCTCTGCCCAGGTGTCTGGGCTGTTGGCCCTGCGCGCCAGGTCCTCCCTGCG 180

CGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACCTACTTCGGACT 839
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACCTACTTCGGACT 240

CAGGTTCACTGAGGGGCCCTGGTCTCCCCACAGTCGTCCCAGGCTGCGCGGCTCCCC 899
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAGGTTCACTGAGGGGCCCTGGTCTCCCCACAGTCGTCCCAGGCTGCGCGGCTCCCC 300

CAGCTCTCTGGGCACCCGGTCCCCTGCCCTCTGCCCACCCTACGCCGCTCTTTGCTCCAG 959
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAGCTCTCTGGGCACCCGGTCCCCTGCCCTCTGCCCACCCTACGCCGCTCTTTGCTCCAG 360

GCCCCTCCCTCTAGAGGCTGCCTGGCCTGTTCACGTGTTTTCATCCCACATAAA 1019
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCCCCTCCCTCTAGAGGCTGCCTGGCCTGTTCACGTGTTTTCATCCCACATAAA 420

GTATTCCCACTCTTATCTTACAACTCCCCCACCCGCCCACTCTCCACCTCACTAGCT 1079
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTATTCCCACTCTTATCTTACAACTCCCCCACCCGCCCACTCTCCACCTCACTAGCT 480

AATCCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCCCCCCTGGCCACAGACCC 1139
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AATCCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCCCCCCTGGCCACAGACCC 540

GGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCAAGAGACCCCACTTCA 1199
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCAAGAGACCCCACTTCA 600

CTAAGAGGGGCTGGACCTGGCGGCAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGA 1259

Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      GGCACCTAAGAGAGGGGCTGACCTGGCGGCAGGAAGCAGACAAGAGACTCCTCCCTTGAGAATTCCTC

Qy 1260 GTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGAGCAAGCTCCCCTTGAGAATTCCTC
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGAGCAAGCTCCCCTTGAGAATTCCTC

Qy 1320 TTTTAAAACAGATATATTTTTATTATTATTGTGACAAAATGTTGATAAATGG 1
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 TTTTAAAACAGATATTATTTTTTATTATTGTGACAAAATGTTGATAAATGG

RESULT 2
BI766766
LOCUS   BI766766
DEFINITION   6O305686F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:520€
             mRNA sequence.   EST 25
ACCESSION   BI766766
VERSION   BI766766.1 GI:15758344
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Butel€
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 834)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection
  JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information ca
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLAM11517 row: c   column: 18
             High quality sequence stop: 772.
FEATURES   Location/Qualifiers
     source   1..834
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5206217"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_122"
                     /note="Organ: pooled lung and spleen; Vector: pCl
                     Site 1: NotI; Site 2: EcoRV (destroyed); RNA sou
                     anonymous pool of 24 week female lung, 16 week f€
                     spleen, and 20-22 week male spleens. Library is
                     primed and directionally cloned (EcoRV site is d€
                     upon cloning). Average insert size 1.4 kb, inse
                     range 1-3 kb. Library is normalized and enriched
                     full-length clones and was constructed by C. Gr
                     (Invitrogen). Research Genetics tracking code (
                     this is a NIH_MGC Library."

ORIGIN
       Query Match   55.8%;   Score 765.8; DB 12;   Length 834;
       Best Local Similarity 98.0%;   Pred. No. 7.3e-134;
       Matches 818;   Conservative   0;   Mismatches   12;   Indels   5;   C

Qy   272 CCCAGGAGGAGCTGGTGCAGAGGAGGACCAGGACCGTCGGAACTGAATCCCCAG
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CCCAGGAGGAGCTGGTGCAGAGGAGGACCAGGACCCGTCGGAACTGAATCCCCAG

Qy   332 AAGAAAGCCAGGATCCTGGCCCTTTCCTGAACCGACTAGTTCGGCCTCGCAGAAGT
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AAGAAAGCCAGGATCCTGCGCCTTTCTCGAACCGACTAGTTCGGCCCTCGCAGAAGT

Qy   392 CTAAAGGCGGGGAAAACACGGGCTCGAGAGCGATCGCAGCCCATTATGAAGTTCA
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
High quality sequence start: 24
High quality sequence stop: 550.

FEATURES             Location/Qualifiers
     source          1..948
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6278608"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC 113"
                     /note="Organ: spleen; Vector: pOTB7; Site_1: XhoI
                     EcoRI; cDNA made by oligo-dT priming. Directional
                     into EcoRI/XhoI sites using the following 5' adap
                     GGCACGAG(G). Library constructed by Ling Hong in
                     laboratory of Gerald M. Rubin (University of Cal
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagen
                     Superscript II RT (Life Technologies). Note: this
                     NIH_MGC Library."

ORIGIN
Query Match        54.9%;    Score 753.2;  DB 13;   Length 948;
Best Local Similarity  95.6%;  Pred. No. 1.8e-131;
Matches 807;  Conservative  0;  Mismatches  30;  Indels  7;   G
```

```
Qy  305 ACCCGTCGGAACTGAATCCCCAGACAGAAGAAAGCCAGGATCCTCGCGCCTTTCCTG.
        ||||||||||||||||||||||||||| |||| ||||||||||||||||||||||||
Db    1 ACCCGTCGGAACTGAATCCCCAGACAGAGAACGCCAGGATCCTCGCGCCTTTCCTG.

Qy  365 GACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGAGAAAACACGGGCTCGAAGA
        |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   61 GACTAGTTCGGCCTCGAAGAGTGCACCTAAAGGCCGAGAAAACACGGGCTCGAAGA

Qy  425 TCGCAGCCATTATGAAGTTCATCCACGACGACCTGACACGAGACGGAGCGCAGGGT
        ||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db  121 TCGCAGCCATTATGAAGTTCATCCACGACGACCTGACACGAGACGGAGCGAGC----GCAGGT

Qy  485 ACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAGCCCCTCTGCGGC
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  177 ACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAGCCCCTCTGCGC

Qy  545 ACCGCCAGATCGGGAGTTTATAGTCACCCGGCTGGGCTTCACTACTGTACTGT
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  237 ACCGCCAGATCGGGAGTTTATAGTCACCCGGCTGGGCTTCACTACTGTACTGT

Qy  605 TGCACTTTGATGAGGGAAGGCTGCTCTACCTGAAGCTGGACTTGCTGGTGATGGT
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  297 TGCACTTTGATGAGGGAAGGCTGCTCTACCTGAAGCTGGACTTGCTGGTGATGGT

Qy  665 TGGCCCTGCGCTGCTGGAGGAATTCTCAGCCACTGCGGCCAGTTCCTCGGGCCC
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  357 TGGCCCTGCGCTGCTGGAGGAATTCTCAGCCACTGCGGCCAGTTCCTCGGGCCC

Qy  725 TCCGCCCTGCCAGGTGTCTGGGCTGTTGGCCCTGCGGG
        ||||||||||||||||||||||||||||||||||||||
Db  417 TCCGCCCTGCCAGGTGTCTGGGCTGTTGGCCCTGCGGG

Qy  785 GCACCCTCCCCATCTCAAGGCTGCCCCCTCCTCACCTACTTCGGACTC
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  477 GCACCCTCCCCATCTCAAGGCTGCCCCCTCCTCACCTACTTCGGACTC

Qy  845 AGGTTCACTGAGGGGCCCGGCCCCCACCAGTCGTCCGGGCTCCCCT
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  537 AGGTTCACTGAGGGGCCCGGCCCCCACCAGTCGTCCGGGCTCCCCT

Qy  905 AGCTCTTCTGGGGCACCCGGTCCCCTCTGCCCCCACCCTCAGCGCGTCTTTGCTCAGA-
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  597 AGCTCTTCTGGGGCACCCGGTCCCCTCTGCCCCCACCCTCAGCGCGTCTTTGCTCCAGA(

Qy  964 GCCCCTCCCTCTAGAGGCTGCTGGGCCTGTTCACGTGTTTCCATCCCACATAAA
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  657 GCCCCTCCCTCTANAGGCTGCTGGGCCTGTTCACGTGTTTCCATCCCCTTAAA

Qy 1024 GTATTCCCACTCTTATCTTACACAACTCCCCCACCCCACTCCACCTCCACCTCACTAGCTC
```

```
5
RT_8353983 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6278608
A sequence.                  948 bp    mRNA    linear    EST 16-JUL-2002
5
5.1   GI:21846084

piens (human)
piens
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 948)
    http://mgc.nci.nih.gov/,
1 Institutes of Health, Mammalian Gene Collection (MGC)
shed (1999)
: Robert Strausberg, Ph.D.
  cgapbs-r@mail.nih.gov
Procurement: Dr. Mark Watson
ibrary Preparation: Rubin Laboratory
library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
quencing by: Agencourt Bioscience Corporation
distribution: MGC clone distribution information can be
hrough the I.M.A.G.E. Consortium/LLNL at:
  image.llnl.gov
LLCM2466  row: n  column: 17
```

```
Qy   549  CCAGATCGGGAGTTTATAGTCACCGGGCTCGGGCTCTCACTACCTGTACTGTCA
               ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   180  CCAGATCGGGAGTTTATAGTCACCGGGCTCGGGCTCTTACTACCTGTACTGTCA

Qy   609  CTTTGATGAGGGAAGGCTGTCTTACCTGAAGTGGACTTGCTGGTGGATGGTGT
               ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   240  CTTTGATGAGGGAAGGCTGTCTTACCTGAAGTGGACTTGCTGGTGGATGGTGT

Qy   669  CCTGCGCTGCTGGAGGAATTCTCAGCCACTCGCGGCCACTCCCTCGGGCCCCA
               ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   300  CCTGCGCTGCTGGAGGAATTCTCAGCCACTCGCGGCCACTCCCTCGGGCCCCA

Qy   729  CCTCTGCCAGTGTCTGGCGTGTTGGCCGTGCCCAGGGTCCTCCCTGCGGAT
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   360  CCTCTGCCAGTGTCTGGCGTGTTGGCCGTGCCCAGGGTCCTCCCTGCGGAT

Qy   789  CCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACCTACTTCGGACTCTT
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   420  CCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACCTACTTCGGACTCTT

Qy   849  TCACTGAGGGGCCCTGGTCTCCCCACAGTGTCCCAGGCTGCGCGGCTCCCCTCGA
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   480  TCACTGAGGGGCCCTGGTCTCCCCACAGTGTCCCAGGCTGCGCGGCTCCCCTCGA

Qy   909  CTCTGGGCACCCGGTCCCCTCGCCCCCTGCGCCCACCCTCAGCCGCGTCTTTGCTCCAGACCT
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   540  CTCTGGGCACCCGGTCCCCTCGCCCCCTGCGCCCACCCTCAGCCGCGTCTTTGCTCCAGACCT

Qy   969  TCCCCTCTAGAGGCTGCTGCCTGGGCCTGTTCACGTGTTTTCATCCACATAAATACA
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   600  TCCCCTCTAGAGGCTGCTGCCTGGGCCTGTTCACGTGTTTTCATCCCCATAAATACA

Qy   1029 CCCACTCTTATCTTACAACTCCCC-ACCGCCACTCTCCACCTACTAGCTCCC
               |||||||||||||||||| ||| | ||||||||||||||||||||||||||
Db   660  CCCACTCCTTATCTTACAACTCCCCCAACCGGCCAACTCTCCAGCTCAATAGCTTCC

Qy   1088 CC----TGACCCTTTGAGGCCCCAGTGATCTCGACTCCCCCTGGCCACAGACC
               ||    ||||||||||||||||||||||||||| |||||||| | ||||||
Db   720  CCCCTGACCCCTTTGAAGGCCCCAGGGATCTCGACTCCCCCTGCCACAAACC

Qy   1144 GGCATT-GTGTTCACTGTACTCTGTGGGCAAGG-ATGGGTCCAGAAGACCCCACT
               |||||| ||||||||||||||||||||||||| |||||||||||||||||||
Db   780  GGCATTGGGGGTTCACTGAACTCTGGGGGAAGGAAGGGGTTCAAAAGACCCCCTT

Qy   1202 CACTAAGAGGGGCT--GGACCTGGGGCAGGAGGCCAAGAGACTGGG--CCTA
               ||||||||||||  |||||||||||||||||||||||||||||  ||||
Db   840  ACTTAAAAGGGGCTTGGGACCTGGGGCAGGAGGAAAGCCCAAAAAAAAATGGGGCCCAAG

Qy   1257 GGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCC  1300
               ||||||||||||||||| |||||||||| || |||| ||
Db   900  GAAATTCCCAAAGGGGAGGGGGGCCGGAAAAAACAAAAAAGTCCCC  943
```

```
       ||||||||||||| |||||||  ||||| |||| ||||||||||||||||| ||||
   TTCCACTCTTTATTTACAAATCCCCCACGCCAATTCTCCACCTCATTAGTTCCCC  776

   CCCTGACCC--TTTGAGGCCCCCAGTGACTCGATCTGCTCCCCCGTCGGCCACGACCCCC  1141
         ||||||  | |||| || |||||  |||||| ||| |||  |||||||||||||||
   CCCTGACCCCTTTGAGGGCGCCCCCAGGGATCTCAACCTCAACCTCCCGTGGGCCCAGAACCCC  836

   CG 1145

   CG 840
```

259
COURT_8303564 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6274716
mRNA sequence.
259
259.1 GI:21782093

sapiens (human)
sapiens
yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
lia; Eutheria; Primates; Catarrhini; Hominidae; Homo
ases 1 to 963)
GC http://mgc.nci.nih.gov/.
nal Institutes of Health, Mammalian Gene Collection (MGC)
lished (1999)
ct: Robert Strausberg, Ph.D.
: cgapbs-r@mail.nih.gov
e Procurement: ATCC
   Library Preparation: Rubin Laboratory
   Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
   Sequencing by: Agencourt Bioscience Corporation
e distribution: MGC clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL at:
//image.llnl.gov
: LLCM2456 row: 1 column: 13
quality sequence stop: 565.
   1..963
   Location/Qualifiers
   1..963
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:6274716"
      /tissue type="epidermoid carcinoma, cell line"
      /lab_host="DH10B (phage-resistant)"
      /clone_lib="NIH_MGC_102"
      /note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;
      Site_2: EcoRI; cDNA made by oligo-dT priming.
      Directionally cloned into EcoRI/XhoI sites using the
      following 5' adaptor: GGCACGAG(G). Library constructed
      by Ling Hong in the laboratory of Gerald M. Rubin
      (University of California, Berkeley) using ZAP-cDNA
      synthesis kit (Stratagene) and Superscript II RT (Life
      Technologies). Note: this is a NIH_MGC Library."

```
larity     54.4%;  Score 747.2;  DB 13;  Length 963;
Conservative    91.9%;  Pred. No. 2.4e-130;
                0;  Mismatches    63;  Indels   13;  Gaps    7;

   TCGGCCCTCGCAGAAGTGCACCTAAGGCCGGAAAACACGGGCTCGAAGAGCGGATCGC  428
   |||||||||||||||||||||||||||||||||||||| |||||||||||||||||
   TCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAA ACGGGCTCGAAGAGCGGATCGC  59

   CCATTATGAAGTTCATCCACGACCTGGACAGGACGGAGCGGCAGGCGTGTGGACGG  488
   |||||||||||||||||||||||||||||||||||||||||||||||||||||
   CCATTATGAAGTTCATCCACGACCTGGACAGGACGGAGCGGCAGGCGTGTGGACGG  119

   AGTGGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCGCGTTACAACCG  548
   |||||||||||||||||||||||||||||||||||||||||||||||
   AGTGGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCGCGCTACAACCG  179
```

RESULT 5
BI819200
LOCUS   BI819200   777 bp   mRNA   linear   EST 0·
DEFINITION   603034614F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:517!
             mRNA sequence.
ACCESSION   BI819200
VERSION   BI819200.1 GI:15930750
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
REFERENCE   1 (bases 1 to 777)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/,
TITLE   National Institutes of Health, Mammalian Gene Collection
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: Life Technologies, Inc.

Library Preparation: Life Technologies, Inc.
Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
sequencing by: Incyte Genomics, Inc.
distribution: MGC clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL at:
/image.llnl.gov
LLAM11437 row: 1 column: 03
uality sequence stop: 759.
1 ..777
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /clone="IMAGE:5175698"
   /lab_host="DH10B"
   /clone_lib="NIH_MGC 115"
   /note="Organ: pooled brain, lung, testis; Vector:
   pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
   source anonymous pool of 6 male brains, age range 23-27; 1
   male lung, age 27; and 1 male testis, age 69. Library is
   oligo-dT primed and directionally cloned (EcoRV site is
   destroyed upon cloning). Average insert size 1.8 kb,
   insert size range 1-3 kb. Library is normalized and
   enriched for full-length clones and was constructed by C.
   Gruber (Invitrogen). Research Genetics tracking code
   021. Note: this is a NIH_MGC Library."

arity   54.1%;   Score 743.4; DB 12; Length 777;
         99.1%;   Pred. No. 1.2e-129;
onservative   0; Mismatches 6;   Indels 1; Gaps 1;

AGGAGGACCAGGACCCGTCGGAACTGAATCCCCAGCACAGAGAAAGCCAGGATCC 347
||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGGAGGACCAGGACCCGTCGGAACTGAATCCCACAGCACAGAAGAAAGCCAGGATCC 72

CCTTTCCTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAAC 407
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCTTTCCTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAAC 132

GCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACGG 467
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTCGACAGGACGG 192

CAGGCAGGTGTGGACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTC 527
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAGGCAGGTGTGGACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTC 252

CCTCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGCTGGGCTCTA 587
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCTCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGCTGGGCTCTA 312

TTGTACTGTCAGGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGGACTT 647
||||||||||||||||||||||||||||||||||||||||||||||||||||||
TTGTACTGTCAGGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGGACTT 372

TTGGATGGTGTGCTGCTGCCCTGCCCTGCCCTGCCTGAGGAATTCTCAGCCACTGCGGCCAG 707
||||||||||||||||||||||||||||||||||||||||||||||||||||||
TTGGATGGTGTGCTGCTGCCCTGCCCTGCCCTGCCTGAGGAATTCTCAGCCACTGCGGCCAGG 432

TTCGGGCCCCAGCTCCGCCTTCGGCCAGGTGTCTGGGCTGTTGGCCCTGCGGCCAGG 767
||||||||||||||||||||||||||||||||||||||||||||||||||||||
TTCGGGCCCCAGCTCCGCCTCGGCCAGGTGTCTGGGCTGTTGGCCCTGCGGCCAGG 492

CCCCTGCGACTCCGGATCGGCACCCTCCCCTGGCCCATCTCAAGGCTGCCCCCTTCCTCAC 827
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCCTGCGACTCCGGATCGGCACCCTCCCCTGGCCCATCTCAAGGCTGCCCCCTTCCTCAC 552

TCGGACTCTTCAGTTCACTGAGGGGCCCTGGTCTCCCACAGTCGTCCAGGC 887
||||||||||||||||||||||||||||||||||||||||||||||||||||||
TTCGGACTCTTCAGTTCACTGAGGGGCCCTGGTCTCCCCAGTCGTCCAGGC 612

GCTCCCCTCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGCTGGGCTCTA 947

Db   613 TGCCGGCTCCCCTGCAGCTCTCTGGGCACCCGGTCCCCTCGCCCACCCTCAC
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   948 TCTTTGTCTCCAGACCTGCCCCTCCCCTCTGAGGCTGCTGGGCCTGTTCACGTGTT
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   673 TCTTTGTCCAGACCTGCCCCTCCCCTCTGAGAGCGCTGCCTGCCTGTTCACGTGTT
Qy   1008 ATCCCACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCC 1052
              ||||||||||||||||||||||||||||||||||||||||||
Db   733 ATCCCACAT-AATACAGTATTCCCACTCTTATCTTCAAATCCCC 776

RESULT 6
BM921213
LOCUS   BM921213   1071 bp   mRNA   linear   EST 12
DEFINITION AGENCOURT_6633046 NIH_MGC_115 Homo sapiens cDNA clone IMAG
            5', mRNA sequence.
ACCESSION   BM921213
VERSION   BM921213.1   GI:19371592
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1071)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (
Unpublished (1999)
REFERENCE
AUTHORS   Contact: Robert Strausberg, Ph.D.
TITLE   Email: cgapbs-r@mail.nih.gov
JOURNAL   Tissue Procurement: Life Technologies, Inc.
COMMENT   cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information ca
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12786 row: p column: 02
            High quality sequence stop: 656.
FEATURES   Location/Qualifiers
     source   1..1071
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5752561"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC 115"
                     /note="Organ: pooled brain, lung, testis; Vector:
                     pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroy
                     source anonymous pool of 6 male brains, age range
                     male lung, age 27; and 1 male testis, age 69. Li
                     oligo-dT primed and directionally cloned (EcoRV s
                     destroyed upon cloning). Average insert size 1.
                     insert size range 1-3 kb. Library is normalized a
                     enriched for full-length clones and was construc
                     Gruber (Invitrogen). Research Genetics tracking
                     021. Note: this is a NIH_MGC Library."

ORIGIN

Query Match   52.8%;   Score 725.4; DB 12; Length 1071;
Best Local Similarity 93.4%;   Pred. No. 3.1e-126;
Matches   836; Conservative   0; Mismatches 46; Indels 13; G

Qy   412 GCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACGG
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12 GCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACGG

Qy   472 CAGGCAGGTGTGGACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTC
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   72 CAGGCAGGTGTGGACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTC

Qy   532 CCTCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGCTGGGCTCTA(
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||

Clone distribution: MGC clone distribution information c;
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12034 row: b column: 07
High quality sequence stop: 728.

FEATURES   Location/Qualifiers
  source   1..731
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:5405478"
              /tissue_type="adenocarcinoma, cell line"
              /lab_host="DH10B (phage-resist)"
              /clone_lib="NIH_MGC_90"
              /note="Organ: liver; Vector: pCMV-SPORT6; Site 1
              Site_2: SalI; Cloned unidirectionally; oligo-dT I
              Average insert size 1.7 kb. Library enriched for
              full-length clones and constructed by Life Techno
              Note: this is a NIH_MGC Library."

ORIGIN

Query Match   51.1%;   Score 701.4; DB 12;   Length 731;
Best Local Similarity   99.3%;   Pred. No. 9.4e-122;
Matches 726; Conservative   0; Mismatches   1; Indels   4;   (

```
Qy   192  CCTGGCGCTGGCCTGCCTCGGCCTCCTCGTCGGCCGTGGTCAGTTTGGGGAGCCGG(
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  CCTGGCGCTGGCCTGCCTCGGCCTCCTCGTCGGCCGTGGTCAGTTTGGGGAGCCGG(

Qy   252  GCTGTCCGCC---CAGGAGCGCTGCCCAGGAGGAGCGCTGGTGGCAGAGGAGACCAG(
          ||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db    61  GCTGTCCGCCCAGCCAGGAGCCTGCCCAGGAGGAGCGCTGGTGGCAGAGGAGGACCAG(

Qy   309  GTCGGAACTGAATCCCAGACAGAAGAAGAGCCAGGATCCTGCGCCTTTCCTGAACC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  GTCGGAACTGAATCCCAGACAGAAGAAGAGCCAGGATCCTGCGCCTTTCCTGAACC

Qy   369  AGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAACACGGGCTCGAAGAGAGCG
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  AGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAACACGGGCTCGAAGAGAGCG

Qy   429  AGCCCATTATGAAGTTCATCCACGACCTGGACAGGACGGAGCGGCAGGCAGGTGTGC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  AGCCCATTATGAAGTTCATCCACGACCTGGACAGGACGGAGCGGCAGGCAGGTGTGC

Qy   489  GACAGTGAGTGGCTGGGAGGAGCCAGAATCAACAGTCCAGCCCTCTGCGCTACC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  GACAGTGAGTGGCTGGGAGGAGCCAGAATCAACAGTCCAGCCCTCTGCGCTAC)

Qy   549  CCAGATCGGGGAGTTATAGTCACCCGGCTGGGCTCTACTGTACTGTCAGC
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  CCAGATCGGGGAGTTATAGTCACCCGGCTGGGCTCTACTGTACTGTCAGC

Qy   609  CTTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTGCTGTGGATGGTGTGC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  CTTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTGCTGTGGATGGTGTGC

Qy   669  CCTGCGCTGCCTGGAGAATTCTCAGCCACTGCGGCCACTGCGCGCGCCTGCGGCCCAGC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  CCTGCGCTGCCTGGAGAATTCTCAGCCACTGCGGCCACTGCGCGCGCCTGCGGCCCAGC

Qy   729  CCCTGCCAGGTGTCTGGCCTGTGTTGGGCCCTGGGCCTCCTCGGAGGTCCTCCCTGGGATCC
          ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   541  CCCTCTGCCAGGTGTCTGGCCTGTGTTGGGCCCTGGGCCTCCTCGGAGGTCCTCCCTGGGATCC

Qy   789  CCTCCCCTGGGCCCCATCCAAGGCTGCCCCCTTCCTCACCTACTTCGGACTCTTCC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   600  CCTCCCCTGGGCCCCATCCAAGGCTGCCCCCTTCCTCACCTACTTCGGACTCTTCC

Qy   849  TCACTGAGGGGCCCTGCTCTCCCCACAGTCGTCGGGCTGCCCGGCTCCCCTGAC
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db   660  TCACTGAGGGGCCCTGCTCTCCCCAGTCGTCGTCCCAGCCAGGCTGCCCGGCTCCCCTGAC
```

TGCGCTACACCGCCAGGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTACTAC 191

ACTGTCAGTGCACTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTG 651
ACTGTCAGTGCACTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTG 251

ATGGTGTGCTGGCCGCTGCCTGGAGAATTCTCAGCCACTGCGGCCCAGTTCC 711
ATGGTGTGCTGGCCCTGCCTGGGAGAATTCTCAGCCACTGCGGCGCAGAGTTCC 311

GGGCCCCAGTCCGGCCTCTGCCAGTGTCTGGGCTGTTGGCCCTGCGGCCCAGGGTCC 771
GGGCCCCAGTCCGGCCTCTGCCAGTGTCTGGGCTGTTGGCCCTGCGGCCCAGGGTCC 371

TGCGGATCCGCACCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACCTAC 831
TGCGGATCCGCACCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACCTAC 431

GACTCTTCCAGGTTCACTGAGGGGCCCCTGTCTCCCCACAGTCGTCCCAGGCTGCC 891
GACTCTTCCAGGTTCACTGAGGGGCCCCTGTCTCCCCACAGTCGTCCCAGGCTGCC 491

CCCCTCGACAGCTCTCTGGGCACCCGGTCCCCTCGCCCCTCAGCCGCTCTT 951
CCCCTCGACAGCTCTCTGGGCACCCGGTCCCCTCGCCCCACCCTCAGCCGCTCTT 551

CCAGACCTGCCCCCCTCCCTAGAGCTGCTGGGCCTGTTCACGTGTTTTCCATCC 1011
CCAGACCTGCCCCCCTCCCTAGAGCTGCTGGGCCTGTTCACGTGTTTTCCATCC 611

TAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCCACTCTCCACCT 1071
TAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCCACTCTCCACCT 671

AGCTCCCCAATCCTCCTGACCCTTTGAGGCCCCAGTGATCTCGACT-CCCCCTGGC 1130
AGCTCCCCAATCCTCCTGACCCTTTGAGGCCCCAGTGATCTCGACTCCCCCTGGC 731

GACCCCC-AGGGCATTGTG-TTCACTGTACTCTGTGGGCAAGGA--TGGGTCCAGA 1186
GACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGAATGGGTCCAAGA 791

CCCACTTCAGG---CACTAAGAGGGGGCTGGACCTGGGCGCGGAGG--AAGCCAAAG 1239
CCCACTTCGGGCACCTAAAAGGGGCCTGGACCTGGGCGCGGAGGCAAGGAAAGCCAAAAG 851

TGGG-CCTAGGCCAGGAGTTCCCAAATGTGAGGGGCCGAGAAACAAGACAAG 1293
TGGGCCCTAGGCCCAGGAGGTCCCCAAATGGGGAGGGGGCCCAGAAAACCAAG 906

11   731 bp   mRNA   linear   EST 11-OCT-2001
825F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5405478 5',
equence.
11
11.1 GI:16045386

apiens (human)
apiens
ota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 731)
C http://mgc.nci.nih.gov/.
al Institutes of Health, Mammalian Gene Collection (MGC)
ished (1999)
t: Robert Strausberg, Ph.D.
    cgapbs-r@mail.nih.gov
  Procurement: ATCC
Library Preparation: Life Technologies, Inc.
Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
equencing by: Incyte Genomics, Inc.

```
       GGCACC 919
       ||||||
       GGCACC 730
```

Qy   552 GATCGGGGAGTTTATAGTCACCCGGCTGGGCTCTACTACCTGTACTGTCAGGTGC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   283 GATCGGGGAGTTTATAGTCACCCGGCTGGGCTCTACTACCTGTACTGTCAGGTGC

Qy   612 TGATGAGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTGTGGATGGTGTGCTGG
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   343 TGATGAGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTGTGGATGGTGTGCTGG

Qy   672 GCGCTGCCTGGAGAATTCTCAGCCACTGCGCGCCAGTTCCTGCGGCCCCAGCTCC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   403 GCGCTGCCTGGAGAATTCTCAGCCACTGCGCGCCAGTTCCTGCGGCCCCAGCTCC

Qy   732 CTGCCAGGTGTCTGGGCTGTTGGCCTGCGGCCCAGGTCTCCCTGCGGATCCGCA
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   463 CTGCCAGGTGTCTGGGCTGTTGGCCTGCGGCCCAGGTCTCCCTGCGGATCCGCA

Qy   792 CCCCTGGCGCCATCTCAAGCTGCCCCTTCCTCACCTACTTCGGACTCTTCCAGG
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   523 CCCCTGGCGCCATCTCAAGCTGCCCCTTCCTCACCTACTTCGGACTCTTCCAGG

Qy   852 CTGAGGGGCCCTGGTCTCCCCACAGTGTCGTCCCAGGCTGCCGGCTCCCCTCGACAGC
          ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   583 CTGAGGGGCCCTGGTCTCCCCGCAGTGTCCCCAGGCTGCCGGCTCCCCTCGACAGC

Qy   912 TGGGCACCCGGTCCCCTGCCCCACCCTCAGCCGCTCTTTGCTCCAGACCTGCCC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   643 TGGGCACCCGGTCCCCTGCCCCACCCTCAGCCACTCTTTGCTCCAGACCTGCCC

Qy   972 CTCTAGAGGCTGCCTGGGCTGTTCACGTGTT---TTCCATCCCACATAAATACA
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   703 CTCTAGAGGCTGCCTGGGCCCTGTTCACAGTTGATATCCACATCCCACATAAATACC

Qy 1026 ATTCCCACTCTTATCTTACAAACTCCCCACCGCCCACTCTCCACCTCA 1073
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   763 TTTCCCACTCTTATCTTACAACTTCCCCCAACAGACCAATCATCATCACA 810

RESULT 9
BI870393
LOCUS   BI870393   728 bp   mRNA   linear   EST 11
DEFINITION   603395641F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:54054!
             mRNA sequence.
ACCESSION   BI870393
VERSION   BI870393.1   GI:1604066
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 728)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (!
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLAM12034   row: a column: 12
             High quality sequence stop: 728.
FEATURES   Location/Qualifiers
     source   1..728
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5405459"
                     /tissue_type="adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"

SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
            1 (bases 1 to 697)
REFERENCE
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE   Normalization and subtraction: two approaches to facilitate
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
  PUBMED   8889548
COMMENT   Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road,   4156   MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9555
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Mary Hendrix
            cDNA Library preparation: Dr. M. Bento Soares, University
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Io
            DNA Sequencing by: Dr. M. Bento Soares, University of Io
            Clone Distribution: Distribution information can be found
            http://genome.uiowa.edu/distribution/humanfl.html
            The following repetitive elements were found in this cDNA
            sequence: 37-143, >GC_rich#Low_complexity
            Seq primer: pYX-5.

FEATURES   Location/Qualifiers
     source   1..697
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30563490"
                     /tissue_type="Chondrosarcoma Lung Metastasis cel
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_MGC_214"
                     /note="Organ: Lung; Vector: pYX-Asc; Site 1: EcoR
                     Site 2: Not I; The library was constructed accord
                     Bonaldo, Lennon and Soares, Genome Research, 6:79
                     1996. Denatured RNA was size fractionated on a 1%
                     gel. First strand cDNA synthesis was primed with
                     primer containing a Not I site. Double strand cDN
                     size selected according to mRNA size fraction, li
                     with EcoR I adaptor, digested with Not I and ther
                     directionally into pYX-Asc vector. The library ta
                     sequence located between the Not I site and the I
                     is GATAAGCCA. Tissue was provided by Mary Hendri

ORIGIN
            Query Match   48.1%; Score 660.4; DB 14; Length 697;
            Best Local Similarity   98.6%; Pred. No. 4.9e-114;
            Matches 688; Conservative   0; Mismatches   6; Indels   4; (

Qy   57 CGGGTCCGGGATGGGGGGGCGGGGTGAGGCAGGCAGCACAGCCCCCGCCCCCCCATGGCCC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CGGGTCCGGGATGGGGGGGCGGGGTGAGGCAGGCAGCACAGCCCCCGCCCCCCCATGGCCC

Qy   117 TCGGAGCCAGAGGCGGAGGGGCGCCGGGGGGAGCCGGGCACCGCCCGCCCTGCTGGTCC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TCGGAGCCAGAGGCGGAGGGGCGCCGGGGGGAGCCGGGCACCGCCCGCCCTGCTGGTCC

Qy   177 CGCGCTGGGCCTGGGCCTGGGCCTGGGCCTGGGCCTGCTGGCCTCCTGGGCTGCTGGCTCA
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CGCGCTGGGCCTGGGCCTGGGCCTGGGCCTGGGCCTGCTGGCCTCCTGCTGGGCTGGTCA

Qy   237 GGGGAGCCGGGCATCGCTGTCCGCC---CAGGAGGGAGGAGAGCTGGTGC
           ||||||||||||||||||||||||| |||||||||||||||||||||||
Db   181 GGGGAGCCGGGCATCGCTGTCCGCCAGCAGGAGGGAGGAGAGCTGGTGC

Qy   294 GGAGGACCAGGACCCGTCGAACTGAATCCCCAGACAGAAGAAAGCCAGGATCCTG
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGAGGACCAGGACCCGTCGAACTGAATCCCCAGACAGAAGAAAGCCAGGATCCTG

Bonaldo, Lennon and Soares, Genome Research, 6:79
1996. Denatured RNA was size fractionated on a 1%
gel. First strand cDNA synthesis was primed with
primer containing a Not I site. Double strand cDN
size selected according to mRNA size fraction, li
with EcoR I adaptor, digested with Not I and then
directionally into pYX-Asc vector. The library ta
sequence located between the Not I site and the p
is GATAAGGCCA. Tissue was provided by Mary Hendri

ORIGIN

Query Match   47.8%; Score 656.6; DB 14; Length 666;
Best Local Similarity 98.9%; Pred. No. 2.5e-113;
Matches 659; Conservative 0; Mismatches 0; Indels 0;   G

```
Qy   292 GAGGAGGACCAGGACCCGTCGGAACTGAATCCCCAGACAGGAGAAAGCCAGATCC
         ||||||||||| ||||||||||||||||||||||||||||| ||||||||||||||
Db     1 GAGGAGGACCCGGACCCGTCGGAACTGAATCCCCAGACAGACAGAAGCCAGATCC

Qy   352 CCTTTCCTGAACCGACTAGTTCGGCCTCGCGAGAAGTGCACCTAAAGGCCGAAAAAC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 CCTTTCCTGAACCGACTAGTTCGGCCTCGCGAGAAGTGCACCTAAAGGCCGAAAAAC

Qy   412 GCTCGAAGAGCAGGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACGG
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GCTCGAAGAGCAGGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACGG

Qy   472 CAGGCAGGTGTGGACGGGACAGTGAGTGGCTGGAGGAAGCCAGAATCAACAGCTC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CAGGCAGGTGTGGACGGGACAGTGAGTGGCTGGAGGAAGCCAGAATCAACAGCTC

Qy   532 CCTCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGCTCTA
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CCTCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGCTCTA

Qy   592 CTGTACTGTCAGGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGGACTT
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CTGTACTGTCAGGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGGACTT

Qy   652 GTGGATGGTGTGCTGGCCCTGCGCTGCCCTGGAGGAATTCTCAGCCACTGCGGCCAG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GTGGATGGTGTGCTGGCCCTGCGCTGCCCTGGAGGAATTCTCAGCCACTGCGGCCAG

Qy   712 CTCGGGCCCCAGCTCGCGCCTCTGCGCCTCGCGCCTCTGCCAGGTGTCTGGCGTGTTGGCCCTGCGGCCAGG
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CTCGGGCCCCAGCTCGCGCCTCTGCGCCTCGCGCCTCTGCCAGGTGTCTGGCGTGTTGGCCCTGCGGCCAGG

Qy   772 TCCCTGCGGATCGCAGCCTCCCCGGGCCATCTCAAGGCTGCCCCTTCCTCAC
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TCCCTGCGGATCGCAGCCTCCCCGGGCCATCTCAAGGCTGCCCCTTCCTCAC

Qy   832 TTCGGACTCTTCCAGGTTCACTGAGGGGCCCTGGTCTCCCCACAGTCGTCCAGGC
         ||||||||||||||||||||||||||||||||||||||  | | ||||||| | ||
Db   541 TTCGGACTCTTCCAGGTTCACTGAGGGGCCCTGGTCTCCTCNCCGCAGTCGTCNCAGCC

Qy   892 GGCTCCCCTCGACAGCTCTTCTCGGGCACCCGGTCCCCTCTGCGCCCACCCTCAGCCGCG
         ||||||| |||||||||| ||||||||||||||| |||||||| ||||| |||||||||
Db   601 GGCTCCCCCTCGACAGCTCTCTCGGGCACCCGGTCNCCTCTGCCCCCACCCCCTCAGCCGCT

Qy   952 TGCTCCC 957
         |||||||
Db   661 TGCTCCC 666
```

RESULT 12
BI966060/c
LOCUS   BI966060   609 bp   mRNA   linear EST 12-
DEFINITION ie72g04.x1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sa
            cDNA clone IMAGE:5672623 3' similar to TR:O54907 O54907 TNF
            WEAK INDUCER OF APOPTOSIS ;, mRNA sequence.
ACCESSION   BI966060
VERSION   BI966060.1   GI:16340465
KEYWORDS   EST.

(continuation / adjacent alignment fragment)

```
TGAACCGACTAGTTCGGCCTCGCGAGAAGTGCACCTAAAGGCCGAAAAACACGGGC 413
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTGAACCGACTAGTTCGGCCTCGCGAGAAGTGCACCTAAAGGCCGAAAAACACGGGC 360

AGAGGCGATCGCAGCCCATTATGAAGCTCTGGCCTGGCTGGACGCGGAGCGCA 473
||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGAGGCGATCGCAGCCCATTATGAAGTTCATCCACGACGACGAGCGCA 420

GGTGGACGGGACAGTGAGTGGCTGGAGGAAGCCAGAATCAACAGCTCCAGCCC 533
|||||||||||||||||||||||||| |||||||||||||||||||||||||
GGTGGACGGGACAGTGAGTGGCTGGAGGAAGCCAGAATCAACAGCTCCAGCCC 480

GCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGCTCTACTACCT 593
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGCTCTACTACCT 540

GTGTCAGGTGCACTTTGATGAGGGGAAGGCTGTCTACCTGCTGACTTGCTGGT 653
||||||||||||||||||||||||||| ||||||||||||||||||||||||||
GTCAGGTGCACTTTGATGAGGGGAAGGCTGTCTACCTGCTGACTTGCTGGT 600

GTGTCTGGCCCTGCGCTGCCCTGCTGGAATTCTCAGCCACCTGCGGCCAGTTCCCT 713
|||||||||||||||||||||||||||||||||||||||| |||||||||||
GTGTCTGGCCCTGCGCTGCCTGGAGG-ATTCTCAGCCACCTGCGGCGAGTTCCCT 659

CCCAGCTCCGCTCTGCTGCCAGGTGTCTGGGCTGT 751
|||||||||||||||||||||||||||||
CCCCAGCTCCGCCCTCTGCGCCAGGTGTCTGGGCTGT 697
```

(adjacent GenBank record, left margin)

```
9                    666 bp    mRNA      linear  EST 05-AUG-2003
TO-avw-m-22-0-UI.r1 NIH_MGC_214 Homo sapiens cDNA clone
0559317 5', mRNA sequence.
9.1   GI:33203878
piens (human)
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 666)
,M.F., Lennon,G. and Soares,M.B.
zation and subtraction: two approaches to facilitate gene
ry
Res. 6 (9), 791-806 (1996)
7
: Soares, MB
ated Laboratory for Computational Genomics
ity of Iowa
ton Road, 4156 MEBRF, Iowa City, IA 52242, USA
9 335 8250
9 335 9565
bento-soares@uiowa.edu
Procurement: Mary Hendrix
ibrary preparation: Dr. M. Bento Soares, University of Iowa
ibrary Arrayed by: Dr. M. Bento Soares, University of Iowa
quencing by: Dr. M. Bento Soares, University of Iowa
Distribution: Distribution information can be found at
genome.uiowa.edu/distribution/humanfl.html
mer: pYX-5.
Location/Qualifiers
1..666
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30559317"
/tissue_type="Chondrosarcoma Lung Metastasis cell lines"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_MGC_214"
/note="Organ: Lung; Vector: pYX-Asc; Site 1: EcoR I;
Site_2: Not I; The library was constructed according
```

Qy 1027 TTCCCACTCTTATCTTACACAACTCCCCACCGCCGCCACTCTCCACCTCACTAGCTCC
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   369 TTCCCACTCTTATCTTACACACTCCCCACCGCCGCCACTCTCCACCTCACTAGCTCC

Qy 1087 CCCTGACCCTTTGAGGCCCCAGTGATCTCGACTCCCCCTGGCCACAGACCCCA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   309 CCCTGACCCTTTGAGGCCCCAGTGATCTCGACTCCCCCTGGCCACAGACCCCA

Qy 1147 ATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCAAGAGACCCCACTTCAGG
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   249 ATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCAAGAGACCCCACTTCAGG

Qy 1207 AGAGGGGTGGACCTGGGCGCAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGAGT
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   189 AGAGGGGTGGACCTGGGCGCAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGAGT

Qy 1267 AATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCGTGGGATTT
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   129 AATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCGTGGGATTT

Qy 1327 AACAGATATTATTTTTATTATTATTGTGACAAAATGTTGATAAATGG 1373
                |||||||||||||||||||||||||||||||||||||||||||||||
Db   69 AACAGATATTATTTTTATTATTATTGTGACAAAATGTTGATAAATGG 23

RESULT 13
AF163779
LOCUS   AF163779   1027 bp   DNA   linear   GSS 2
DEFINITION   AF163779 Human Homo sapiens genomic clone BAC750E14, genor
             sequence.
ACCESSION   AF163779
VERSION   AF163779.1 GI:5726439
KEYWORDS   GSS.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
            1 (bases 1 to 1027)
REFERENCE   Cousin,P., Billotte,J., Chaubert,P. and Shaw,P.H.
AUTHORS   Physical map of 17p13 and the genes adjacent to p53
TITLE   Genomics 63 (1), 60-68 (2000)
JOURNAL   20130114
MEDLINE   10662545
PUBMED   Contact: Shaw PH
COMMENT   Experimental Oncology
            Institute of Pathology
            Rue du Bugnon 25, Lausanne, VD 1011, Switzerland
            sub_clone=AB2R Asc-BamHI PSL1180
            Class: BAC subclone.
FEATURES   Location/Qualifiers
     source   1..1027
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /map="17p"
                     /clone="BAC750E14"
                     /clone_lib="Human"

ORIGIN

Query Match   42.3%; Score 581.4; DB 28; Length 1027;
Best Local Similarity   98.0%; Pred. No. 4.1e-99;
Matches 577; Conservative   8; Mismatches   4; Indels   0; (

Qy   785 GCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACCTACTTCGGACT
               ||||||||||||||||||||||||| |||||||||::::|||||||||||||||||
Db   1 GCWCCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACCTACTTCGGACT

Qy   845 AGGTTCACTGAGGGCGCCTGGTCTCCCACCAGTCGTCCCAGGCTGCCGGCTCCCC
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AGGTTCACTGAGGGCGCCTGGTCTCCCACCAGTCGTCCCAGGCTGCCGGCTCCCC

Qy   905 AGCTCTCTCGGCACCCGGTCCCCTCTGCCCCACCCTCAGCCGCTCTTTGCTCCAG

(Un-normalized, unamplified): cs"
/note="Organ: Eye; Vector: pCMVSPORT6; Two differ...
eyes (75-80 years old) yielded approximatly 600 ...
dissected RPE/choroid tissue. This in turn yield...
of total RNA and 7 ug of mRNA. A directionally cl...
library in the pCMVSPORT6 vector was constructed...
Technologies (Rockville, MD; now part of Invitrog...
essentially following the protocols of the SuperS...
Plasmid System (Invitrogen Corp.
<http://www.invitrogen.com/>). The library code
designation was cs. For this library, cDNA insert...
cloned into the NotI/MluI sites of the vector. ES...
analysis was performed on the unamplified library
NIH Intramural Sequencing Center (NISC)."

ORIGIN

Query Match   41.3%;   Score 567;   DB 14;   Length 567;
Best Local Similarity 100.0%;   Pred. No. 1.8e-96;
Matches 567; Conservative 0; Mismatches 0; Indels 0; G...

```
Qy    55  CTCGGTCCCGGATGGGGCGGGGGTGAGGCAGGCAGCCCCCGCCCCCATGGC...
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  CTCGGTCCCGGATGGGGCGGGGGTGAGGCAGGCAGCCCCCGCCCCCATGGC...

Qy   115  CGTCGGAGCCAGAGGCGGAGGGGGCGCCCGGGGAGCCCGGGCACCGCCCTGCTGGT...
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  CGTCGGAGCCAGAGGCGGAGGGGGCGCCCGGGGAGCCCGGGCACCGCCCTGCTGGT...

Qy   175  CTCGCGCTGGGCCTGGGCCTGCGCCTGCGCCTCTCGGCCTCTCGCTGCTGGCCGTGGT...
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  CTCGCGCTGGGCCTGGGCCTGCGCCTGCGCCTCTCGGCCTCTCGCTGCTGGCCGTGGT...

Qy   235  TTGGGGAGCCGGGCATCGCTGCTGCCGCCCAGGAGCCTGCCCAGGAGGAGCTGGTGGCG...
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  TTGGGGAGCCGGGCATCGCTGCTGCCGCCCAGGAGCCTGCCCAGGAGGAGCTGGTGGCG...

Qy   295  GAGGACCAGGACCCGTCGGAACTGAATCCCAGACAGAAGCCAGGATCCTGC...
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  GAGGACCAGGACCCGTCGGAACTGAATCCCAGACAGAAGCCAGGATCCTGC...

Qy   355  TTTCTGAACCGACTAGTTCGGCCTTCGCAGAAGTGCACCTAAAGGCCGGAAAACACG...
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  TTTCTGAACCGACTAGTTCGGCCTTCGCAGAAGTGCACCTAAAGGCCGGAAAACACG...

Qy   415  CGAAGAGGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGACACAGGACGAGGC...
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  CGAAGAGGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGACACAGGACGAGGC...

Qy   475  GCAGGTGTGGACGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAG...
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  GCAGGTGTGGACGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAG...

Qy   535  CTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTACTA...
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  CTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTACTA...

Qy   595  TACTGTCAGGTGCACTTTGATGAGGGG  621
          |||||||||||||||||||||||||||
Db   541  TACTGTCAGGTGCACTTTGATGAGGGG  567
```

RESULT 15
BQ884231
LOCUS   BQ884231   AGENCOURT_8682031 Lupski_sciatic_nerve Homo sapiens cDNA cl...
                        940 bp   mRNA   linear   EST 16-
DEFINITION IMAGE:6197488 5', mRNA sequence.
ACCESSION   BQ884231
VERSION   BQ884231.1 GI:22276239
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelec...
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```
          |||||||||||||||||||||||||||||||||:||::|||||||||||||||||||
          TCTGGGCACCCGGTCCCCGCTGYYCCACCGCGCTCTTTGCTCCAGACCTG  180

          CCCTCTAGAGGCTGCCTGGGCCTGTTCACGTGTTTCCATCCCACATAAATACAG  1024
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
          CCCTCTAGAGGCTGCCTGGGCCTGTTCACGTGTTTCCATCCCACATAAATACAG  240

          CCACTCTTATCTTACAACTCCCCACCGGCCACTCTCCACCTCACTAGCTCCCA  1084
          ||||||||||||||||    :|||||||||||||||||||||||||||||||||
          CCACTCTTATCTTACAANNNNYCCACCGCCCACTCTCCACCTCACTAGCTCCCA  300

          TGACCCTTTGAGGCCCCCAGTGATCTCGACTCCCCCCGGCCACGACCCCCAGG  1144
          ||||||||||||||::||||||||||||||||||||||||||||||||||||||
          TGACCCTTTGAGGSCCCCAGTGATCTGACTCCCCCCTGGCCACAGACCCCCAGG  360

          GTGTTCACTGTACTCTGTGTGGGCCAAGGATGGGTCCAGAAGACCCCACTTCAGGCAC  1204
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          GTGTTCACTGTACTCTGTGTGGGCCAAGGATGGGTCCAGAAGACCCCACTTCAGGCAC  420

          GGGGCTGACCTGGCCGGCGGCAGGAAGCCAAAGAGACTGGGCCTCAGGCCAGGAGTTCC  1264
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          GGGGCTGACCTGGCCGGCGGCAGGAAGCCAAAGAGACTGGGCCTCAGGCCAGGAGTTCC  480

          GTGAGGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTT  1324
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          GTGAGGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTT  540

          AGATATTATTTTTATTATTATTGTGACAAAATGTTGATAAATGG  1373
          ||||||||||||||||||||||||||||||||||||||||||||
          AGATATTATTTTTATTATTATTGTGACAAAATGTTGATAAATGG  589
```

9   .y1 Human Retinal pigment epithelium/choroid cDNA
-.malized, unamplified): cs Homo sapiens cDNA clone cs80h07
A sequence.
9
9.1 GI:24733297
piens (human)
piens
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 567)
G., Bernstein,S.L., Wyatt,M.K., Farris,R.N., Behal,A.,
n,J.W., Bouffard,G., Smith,D. and Peterson,K.
ed sequence tag analysis of human RPE/choroid for the
Project: Over 6000 non-redundant transcripts, novel genes
ice variants
s. 8 (4), 205-220 (2002)
0
0
: Wistow G
on Molecular Structure and Function
l Eye Institute
NIH, Bethesda, MD 20892-2740, USA
1 402 3452
1 496 0078
graeme@helix.nih.gov
80 row: h column: 07
mer: M13Rp1 reverse primer (ABI).
           Location/Qualifiers
1..567
           /organism="Homo sapiens"
           /mol_type="mRNA"
           /db_xref="taxon:9606"
           /clone="cs80h07"
           /tissue_type="RPE/choroid"
           /dev_stage="Adult"
           /lab_host="EMDH10B"
           /clone_lib="Human Retinal pigment epithelium/choroid cDNA

ases 1 to 940)
3C http://mgc.nci.nih.gov/.
nal Institutes of Health, Mammalian Gene Collection (MGC)
lished (1999)
ct: Robert Strausberg, Ph.D.
: cgapbs-r@mail.nih.gov
e Procurement: Dr. James R. Lupski
Library Preparation: Life Technologies, Inc.
Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Sequencing by: Agencourt Bioscience Corporation
e distribution:MGC clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL at:
//image.llnl.gov
: LLAM13607 row: j column: 17
quality sequence stop: 453.
1..940
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6197488"
/sex="male"
/tissue_type="sciatic nerve"
/dev_stage="adult, 70 yr"
/lab_host="DH10B"
/clone_lib="Lupski sciatic nerve"
/note="vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCACGGTCCG-3' and
5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.87 kb. This is a primary
library, non-amplified.Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

larity   40.9%;   Score 562; DB 13; Length 940;
96.5%;   Pred. No. 1.8e-95;
Conservative   0; Mismatches 15; Indels   9; Gaps   8;

```
GACCCGTCGGAACTGAATCCCAGACAGAGAGAAAGCCAGAATCCTGCGCCTTCCTG 360
||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
GACCCGTCGGAACTGAATCCCAGACAGAGAGAAAGCCAGGATCCTGCGCCTTCCTG  60

CGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAACACGGGCTCGAAGA 420
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAACACGGGCTCGAAGA 120

ATCGGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACGGAGCGCAGGCAGGT 480
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ATCGGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACGGAGCGCAGGCAGGT 180

GACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAGCCCCTCTGCGC 540
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
GACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAGCCCCTCTGCGC 240

AACCGCCAGATCGGGGAGTTTATAGTCACCCGGGGCTCTCTACTACCTGTACTGT 600
|||||||||||||||||||||||||||||||||||||||||||||||||||||
AACCGCCAGATCGGGGAGTTTATAGTCACCCGGGGCTCTCTACTACCTGTACTGT 300

GTGCACTTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGGATGGT 660
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTGCACTTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGGATGGT 360

CTGGCCCTGCGCGCTGCCTGGAGGAATTCTCAGCCACTGCGCGCCAGTTCCCTCGGGCCCC 720
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTGGCCCTGCGCGCTGCCTGGAGGAATTCTCAGCCACTGCGCGCCAGTTCCCTCGGGCCCC 420

CTCCCCGCCTCTG-CCAGGTGTCTGGGCGTGTTGGCCCTGGCCCCGCCA-GGGTCCTCCCTGC 778
|||||||||||| |||||||| ||||||||||||||||||||||||||
CTCCGCCTCTGCGCCAGGTGTGTCTGGGCGTGTTGGCCCTGGCCCCGCCACTCCCTCCCTGC 480
```

```
Qy  779 GGATCCGCACCCTCCCTGGGCCCATCTCAAGGCTG-CCCCTTCCTCCACCTAC-
        |||||||||||||||||||||||||||||||||||| |||||| |||| |||||
Db  481 GGATCCGCACCCTCCCTGGGCCCATCTCAAGGCTGCCCCCCCTTTCTCACCTACT

Qy  837 ACTCTTCCAGGTTCACTGAGGGGCCCTGGTCTCCCAGCAGTCGT-CCCAGGCTGC
        ||||||||||||||||||||||||||| ||||||||| ||||||| |||||||||
Db  541 ACTCTTCCAGGTTCACTGAGGGGCCCGGGTCTCCCACAGTCGTCCCCAGGCTGC

Qy  895 TCCCCTCGACAGCTCT-GGGCACCGGTCCCTCGCCCACCC--TCAGCCG
        ||||||||||||||||| |||  ||||||||| |||||||| |||||| |
Db  601 TCCCCTCGACAGCTCTCGGGGAACCGGTCCCGTCCCCACCCTGCCCCCACCCCTCAGGCGG

Qy  952 TGCTCCAGACCTGCCCCTCCCTCT 975
        || |||||| ||| ||||||||| |
Db  661 TGGTCCAGAACTGGCCCCTCCCCT 684
```

Search completed: April 7, 2004, 23:15:25
Job time : 3339.48 secs

score greater than or equal to the score of the result being pri
and is derived by analysis of the total score distribution.

SUMMARIES

ic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

ril 7, 2004, 17:30:20 ; Search time 5268.84 Seconds
                         (without alignments)
                         11294.703 Million cell updates/sec

-09-245-198A-3
73
atgtcattgtagacttga.........gacaaaatgtgtgataaatgg 1373

ENTITY_NUC
pop 10.0 , Gapext 1.0

70272 seqs, 21671516995 residues

ts satisfying chosen parameters: 6940544

gth: 0
gth: 2000000000

inimum Match 0%
aximum Match 100%
isting first 45 summaries

enEmbl:*
gb_ba:*
gb_htg:*
gb_in:*
gb_om:*
gb_ov:*
gb_pat:*
gb_ph:*
gb_pl:*
gb_pr:*
gb_ro:*
gb_sts:*
gb_sy:*
gb_un:*
gb_vi:*
em_ba:*
em_fun:*
em_hum:*
em_in:*
em_mu:*
em_om:*
em_or:*
em_pat:*
em_ph:*
em_pl:*
em_ro:*
em_sts:*
em_un:*
em_vi:*
em_htg_hum:*
em_htg_inv:*
em_htg_other:*
em_htg_mus:*
em_htg_pln:*
em_htg_rod:*
em_htg_mam:*
em_htg_vrt:*
em_sy:*
em_htgo_hum:*
em_htgo_mus:*
em_htgo_other:*

the number of results predicted by chance to have a

Result No.	C	Score	% Query Match	Length	DB	ID	Description
1		1373	100.0	1373	6	BD062758	BD062758 A
2		1325	96.5	1421	6	BD090952	BD090952 Ap
3		1320.2	96.2	1353	6	AX201324	AX201324 Se
4		1320.2	96.2	1353	9	AF358870	AF358870 Hc
5		1320.2	96.2	1368	9	AF055872	AF055872 Hc
6		1285	93.6	1306	9	AF030099	AF030099 Hc
7		1226.4	89.3	1236	6	AR140407	AR140407 Se
8		1226.4	89.3	1236	6	BD057124	BD057124 Me
9		1096.8	79.9	1642	9	BC019047	BC019047 Hc
10	C	768.2	56.0	60268	2	AC016876	AC016876 Hc
11		754	54.9	218485	2	AC127470	AC127470 Pa
12		683.4	49.8	1239	10	AF030100	AF030100 M
13		629.2	45.8	898	6	AX180714	AX180714 Se
14		614.6	44.8	1168	6	BD062757	BD062757 A
15		553	40.3	1816	9	AY081051	AY081051 Hc
16		364.2	26.5	148555	2	AC126921	AC126921 Bc
17	C	309.6	22.5	203083	2	AC069459	AC069459 Mu
18	C	309.6	22.5	234182	10	AL603707	AL603707 M
19		302.2	22.0	130254	2	AC136195	AC136195 Ra
20	C	302.2	22.0	165316	2	AC119115	AC119115 Ra
21		302.2	22.0	223877	2	AC098923	AC098923 Ra
22	C	302.2	22.0	225077	2	AC136563	AC136563 Ra
23		283.4	20.6	149736	2	AC126239	AC126239 Fe
24	C	267	19.4	234801	2	AC118309	AC118309 Ra
25		246.4	17.9	180222	2	AC130192	AC130192 Su
26		242.4	17.7	176258	2	AC126925	AC126925 Ca
27		193.4	14.1	195	6	AX379024	AX379024 Se
28		184.2	13.4	212093	2	AC126237	AC126237 Ca
29		87.2	6.4	7218	6	I66494	I66494 Sequ
30		77.6	5.7	125020	9	AF429315	AF429315 Sequ
31	C	74.8	5.4	303091	2	AC084799	AC084799 Mu
32		70.8	5.2	73948	2	AC022556	AC022556 Ho
33		70.6	5.1	144979	2	AC016280	AC016280 Hc
34		70.4	5.1	220469	2	AC074307	AC074307 Mu
35	C	70.2	5.1	298166	2	AC087563	AC087563 Ho
36		69.2	5.0	936	8	CNS01AYK	CNS01AYK Bc
37		69.2	5.0	82897	2	AC022595	AC022595 Hc
38	C	68.4	5.0	202083	2	AC023833	AC023833 Mu
39		68.2	5.0	110737	2	AC011105	AC011105 Hc
40	C	68.2	5.0	265537	2	AC087228	AC087228 Mu
41	C	68.2	5.0	266935	2	AC144767	AC144767 Mu
42		68	5.0	86784	2	AC138362	AC138362 Mu
43		68	5.0	300695	2	AC079431	AC079431 Mu
44	C	67.6	4.9	626	8	CNS018RP	CNS018RP Bc
45	C	67.2	4.9	3163	6	AX647321	AX647321 Se

ALIGNMENTS

RESULT 1
BD062758
LOCUS   BD062758   1373 bp   DNA   linear   PAT 27
DEFINITION A tumor necrosis factor related ligand.
ACCESSION   BD062758
VERSION   BD062758.1 GI:22608361
KEYWORDS   JP 2001505407-A/2.
SOURCE   unidentified
  ORGANISM unidentified
            unclassified.
REFERENCE   1 (bases 1 to 1373)
  AUTHORS   Chicheportiche,Y. and Browning,J.L.
  TITLE   A tumor necrosis factor related ligand
  JOURNAL   Patent: JP 2001505407-A 2 24-APR-2001;
            BIOGEN INC, THE FACULTY OF MEDICINE OF THE UNIVERSITY OF GE

TNF family related protein
JP 2001505407-A/2
24-APR-2001
07-AUG-1997 JP 1998508239
07-AUG-1996 US   60/023541,18-OCT-1996 US   60/028515 PR
R-1997 US   60/040820
/VES CHICHEPORTICHE,JEFFREY L BROWNING
C12N15/28,C07K14/525,G01N33/68,C07K16/24,C12N15/11,A61K48/00,
C12N5/10,
A61K39/395,A61K38/19,C07K14/705,C12N15/12
Strandedness: Double;
Topology: Linear;
Key   Location/Qualifiers
CDS   1..852.

Location/Qualifiers
1..1373
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

arity 100.0%;   Score 1373;   DB 6;   Length 1373;
Conservative 100.0%;   Pred. No.1.8e-241;
Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy   CAGCTCCGCCTCTGCCAGGTGTCTGGGCTGTTGGCCTCGGGGCCAGGGTCCTCCC 720
Db   CAGCTCCGCCTCTGCCAGGTGTCTGGGCTGTTGGCCTCGGGGCCAGGGTCCTCCC 720

RESULT 2
BD090952
LOCUS   BD090952   1421 bp   DNA   linear   PAT 2
DEFINITION   Apo-3 ligand polypeptide.
ACCESSION   BD090952
VERSION   BD090952.1 GI:22636552
KEYWORDS   JP 2001522584-A/1.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutel
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
REFERENCE   1 (bases 1 to 1421)
  AUTHORS   Ashkenazi,A.J., Marsters,S.A. and Pitti,R.
  TITLE   Apo-3 ligand polypeptide
  JOURNAL   Patent: JP 2001522584-A 1 20-NOV-2001;
             GENENTECH INC
COMMENT   OS   Homo sapiens (human)
             PN   JP 2001522584-A/1
             PD   20-NOV-2001
             PF   09-OCT-1998 JP 2000516042
             PR   10-OCT-1997 US   60/062037,17-DEC-1997 US   60/069
             AVI J ASHKENAZI,SCOT A MARSTERS,ROBERT PITTI
             PC   C12N15/09,A61K38/00,C07K14/705,C07K16/24,C12N15/00,A
             CC   Apo-3 ligand polypeptide
FH   Key   Location/Qualifiers
FT   source   1..1421
FT   /organism=Homo sapiens (human)'.
FEATURES   Location/Qualifiers

```
Qy   977  GAGGCTGCCTGGGCCTGTTCACGTGTTTTCCATCCACATAAATACAGTATTCCCA
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   963  GAGGCTGCCTGGGCCTGTTCACGTGTTTTCCATCCACATAAATACAGTATTCCCA

Qy  1037  TATCTTACAACTCCCCCACCCACTCTCCACCTCCCCCTGGCCACAGACCCCAGGCACTCCTGA
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1023  TATCTTACAACTCCCCCACCCACCGCCACTCTCCACCTCACTAGCTCCCAATCCCTGA

Qy  1097  TTGAGGCCCCCAGTGATCTGACTCCCCCCTGGCCACAGACCCCAGGCACTAAGAGGG
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1083  TTGAGGCCCCCAGTGATCTGACTCCCCCCTGGCCACAGACCCCAGGCACTAAGAGGG

Qy  1157  CTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCCACTTCAGGCAGGAGTTCCCAAATGTG
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1143  CTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACTCGGCCAGGAGTTCCCAAATGTG

Qy  1217  GACCTGGCGGCAGGAAGCCAAAGAGACTGGGCCAGGCCAGGAGTTCCCAAATGTG
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1203  GACCTGGCGGCAGGAAGCCAAAGAGACTGGGCCAGGCCAGGAGTTCCCAAATGTG

Qy  1277  GCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGA
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1263  GCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGA

Qy  1337  ATTTTTATTATTATTGTGACAAAATGTTGATAAATGG  1373
               |||||||||||||||||||||||||||||||||||||
Db  1323  ATTTTTATTATTATTGTGACAAAATGTTGATAAATGG  1359
```

RESULT 3
AX201324
LOCUS   AX201324   1353 bp   DNA   linear   PAT 30-
DEFINITION   Sequence 3 from Patent WO0153486.
ACCESSION   AX201324
VERSION   AX201324.1   GI:15391154
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ashkenazi,A.J., Goddard,A., Godowski,P.J., Gurney,A.L.,
                Hillan,K.J., Marsters,S.A., Pan,J., Pitti,R.M., Roy,M.A.,
                Stone,D.M., Watanabe,C.K. and Wood,W.I.
  TITLE   Compositions and methods for the treatment of tumour
  JOURNAL   Patent: WO 0153486-A 3 26-JUL-2001;
                Genentech, Inc. (US)
FEATURES   Location/Qualifiers
     source   1..1353
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match   96.2%; Score 1320.2; DB 6; Length 1353;
Best Local Similarity   99.8%; Pred. No. 8.4e-232;
Matches 1322; Conservative   0; Mismatches   3; Indels   0; G

```
Qy    49  CGATCCCTGCGGTCCCGGGATGGGGGGGCGGCGGTGAGGCAGGCACAGCCCCCGGCACGCGCCCC
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  CGATCCCTGCGGTCCCGGGATGGGGGGGCGGCGGTGAGGCAGGCACAGCCCCCGGCACGCGCCCC

Qy   109  GCCCGCCGTCGGAGCCAGGAGGCGGAGGCGCCCGGGGGAGGGGCGGCCGGGAGCGGCACCGCGCCC
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  GCCCGCCGTCGGAGCCAGGAGGCGGAGGCGCCCGGGGGAGGGGCGGCCGGGAGCGGCACCGCGCCC

Qy   169  GTCCCGCTCGCGCTCGCTGGGCCTGGGCCTGCCTGGGCCTCGGCCTCGGCCTCGGCCTCGCTGCTGGC
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  GTCCCGCTCGCGCTCGCTGGGCCTGGGCCTGCCTGGGCCTCGGCCTCGGCCTCGGCCTCGCTGCTGGC

Qy   229  GTCAGTTTGGGGAGGCCGCCGCGCATCGGCGCCTGCTCGTCGCTGGCCTGCCCAGAGGAGGAGCGTC
```

1..1421
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```
         96.5%;  Score 1325;  DB 6;  Length 1421;
         98.5%;  Pred. No. 1.1e-232;
         0;  Mismatches   20;  Indels   0;  Gaps   0;
```

```
Qy    ATTTCCGCCGCCGGCTCCCCTCCCCCGATCCCTCGGGTCCCGGGATGGGGGGG   76
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    ATCCCTGACCTGACCCACGCGTCCGCGCATCCCTCGGGTCCCGGGATGGGGGGG   62

Qy    AGGCAGGCACAGCCCCCGCCCGCCCATGGCCGGCCGTCGGAGCGCCAGAGGCGGAGGG  136
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    AGGCAGGCACAGCCCCCGCCCGCCCATGGCCGGCCGTCGGAGCGCCAGAGGCGGAGGG  122

Qy    CGGGGGAGCGGCCTCCTGCTGCCGTGTCCCGCTCGCGCTCGGCCTCGGCCTCGG  196
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    CGGGGGAGCGGCCTCCTGCTGCCGTGTCCCGCTCGCGCTCGGCCTCGGCCTCGG  182

Qy    GCCTGCCTGGCCCTGCTGCCGTGTCGTGGCCGTGGTCAGTTTGGGGAGGCCGCCGCGCATCGCTGT  256
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    GCCTGCCTGGCCCTGCTGCCGTGTCGTGGCCGTGGTCAGTTTGGGGAGGCCGCCGCGCATCGCTGT  242

Qy    CAGGAGCAGCAGAAGAAAGCCAGGATCCTGCGCCTTCTCGAACCGACTAGTTCGGC  316
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    CCCCAGAGCAGCAGAAGAAAGCCAGGATCCTGCGCCTTCTCGAACCGACTAGTTCGGC  302

Qy    CCCCAGACAGCAGAAGAAAGCCAGGATCCTGCGCCTTCTCTGAACCGACTAGTTCGGC  376
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    CCCCAGACAGCAGAAGAAAGCCAGGATCCTGCGCCTTCTCTGAACCGACTAGTTCGGC  362

Qy    AGAAGTGCACCTAAAGGCGGAAAACACGGGCTCGAAGAGCGATCGCAGCCCATT  436
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    AGAAGTGCACCTAAAGGCGGAAAACACGGGCTCGAAGAGCGATCGCAGCCCATT  422

Qy    GTTCATCCACGACCTGACAGGACGGACGGAGCGCAGGCAGCGTGTGACGGACAGTGA  496
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    GTTCATCCACGACCTGACAGGACGGACGGAGCGCAGGCAGCGTGTGACGGACAGTGA  482

Qy    TGGGAGGAAGCCAGAATCAACAGCTCCAGCCCCTCCGGCTACAACCGCCAGATCG  556
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    TGGGAGGAAGCCAGAATCAACAGCTCCAGCCCCTCCGGCTACAACCGCCAGATCG  542

Qy    TTTATAGTCACCCGGGCTGGGCTCTACTACCGTGTCAGTGTCACTTTGATG  616
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    TTTATAGTCACCCGGGCTGGGCTCTACTACCGTGTCAGTGTCACTTTGATG  602

Qy    AAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGGATGGTGCTGGCCCTGCGCCT  676
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    AAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGGATGGTGCTGGCCCTGCGCCT  662

Qy    GAGGAATTCTCAGCCACTGCGGCCCAGTTCCCGGGGCCCCCAGCTCCGGCCTCGCC  736
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    GAGGAATTCTCAGCCACTGCGGCCCAGTTCCCGGGGCCCCCAGCTCCGGCCTCGCC  722

Qy    TCTGGGCTGTTGGCCCTGCGCCCCTTCCTCCTGCGATCCGCACCCTCCCCT  796
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    TTCTGGGCTGTTGGCCCTGCGCCCCTTCCTCCTGCGATCCGCACCCTCCCCT  782

Qy    CATCTCAAGGCTGCCCCCTTCCTCCACCTACTTCGGACTTCTTCCAGGTTCACTCAG  856
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    CATCTCAAGGCTGCCCCCTTCCTCCACCTACTTCGGACTTCTTCCAGGTTCACTCAG  842

Qy    CTGGTCTCCCCACAGTCGTCCCAGGCTGCCGGCTCCCCCTGACAGTCTCTGGGC  916
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    CTGGTCTCCCCGCAGTCGTCCCAGGCTGCCGGCTCCCCCTCGACAGTCTCTGGGC  902

Qy    GTCCCCTCTGCCCACCCTCAGCGCCTCTTTGCTTCCAGACCTGCCCTCCCCTCTA  976
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    GTCCCCTCTGCCCACCCTCAGCGCCTCTTTGCTTCCAGACCTGCCCTCCCCTCTA  962
```

```
Qy        1369 AATGG 1373
                    |||||
Db        1321 AATGG 1325
```

```
RESULT 4
AY358870
LOCUS           AY358870                1353 bp    mRNA    linear   PRI 0.
DEFINITION      Homo sapiens clone DNA30879 Apo-3 Ligand (UNQ181) mRNA, c
                cds.
ACCESSION       AY358870
VERSION         AY358870.1  GI:37182857
KEYWORDS        FLI_CDNA.
SOURCE          Homo sapiens (human)
  ORGANISM      Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutel
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
REFERENCE       1  (bases 1 to 1353)
  AUTHORS       Clark,H.F., Gurney,A.L., Abaya,E., Baker,K., Baldwin,D.,
                Chen,J., Chow,B., Chui,C., Crowley,C., Currell,B., Deuel,
                Dowd,P., Eaton,D., Foster,J., Grimaldi,C., Gu,Q., Hass,P.
                Heldens,S., Huang,A., Kim,H.S., Klimowski,L., Jin,Y., Joh
                Lee,J., Lewis,L., Liao,D., Mark,M., Robbie,E., Sanchez,C.
                Schoenfeld,J., Seshagiri,S., Simmons,L., Singh,J., Smith,
                Stinson,J., Vagts,A., Vandlen,R., Watanabe,C., Wieand,D.,
                Xie,M.H., Yansura,D., Yi,S., Yu,G., Yuan,J., Zhang,M., Zh
                Goddard,A., Wood,W.I. and Godowski,P.
  TITLE         The Secreted Protein Discovery Initiative (SPDI), a Large
                Effort to Identify Novel Human Secreted and Transmembrane
                A Bioinformatics Assessment
  JOURNAL       Genome Res. 13 (10), 2265-2270 (2003)
   PUBMED       12975309
REFERENCE       2  (bases 1 to 1353)
  AUTHORS       Clark,H.F.
  TITLE         Direct Submission
  JOURNAL       Submitted (01-AUG-2003) Department of Bioinformatics, Gen
                Inc., 1 DNA Way, South San Francisco, CA 94080, USA
FEATURES             Location/Qualifiers
     source          1..1353
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DNA30879"
     gene            1..1353
                     /locus_tag="UNQ181"
     CDS             58..807
                     /locus_tag="UNQ181"
                     /note="PRO207"
                     /codon_start=1
                     /product="Apo-3 Ligand"
                     /protein_id="AAQ89229.1"
                     /db_xref="GI:37182858"
                     /translation="MAARRSQRRGRRGEPGTALLVPLALGLGLALAC
                     GSRASLSAQEPAQEELVAEEDQPSELNPQTEESQDPAPFLNRLVRPRI
                     ARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYNRQI(
                     YYLYCQVHFDEGKAVVLKLDLLVDGVLALRCLEEFSATAASSLGPQLR!
                     RPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH"
ORIGIN
```

```
Query Match            96.2%;   Score 1320.2;  DB 9;   Length 1353;
Best Local Similarity  99.8%;   Pred. No. 8.4e-232;
Matches 1322;  Conservative  0;  Mismatches  3;  Indels  0;  (
```

```
Qy     49  CGATCCCCTCGGGTCCCGGGATGGGGGGCGGTGAGGCAGGCAGGCACAGCCCCCGGCCC
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  CGATCCCCTCGGGTCCCGGGATGGGGGGCGGTGAGGCAGGCAGGCACAGCCCCCGGCCC

Qy    109  GCCGCCCGTCGGGAGCCAGAGGCCGAGGGGCGCGGGGGAGCGCGGGAGCCACCGGCCCC
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  GCCGCCCGTCGGGAGCCAGAGGCCGAGGGGCGCGGGGGAGCGCGGGAGCCACCGGCCCC

Qy    169  GTCCCGCTCGCGCTGGGCCTGGGCCTGGCGCTGGCCTGCGGCTCCCGGCCTCCTGCTGG
```

Db 1201 CTAGGCCAGGAGTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTG
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1309 TTCCCTGTGGGATTTTTAAAACAGATATATTTTATTATTATTGTGACAAAATGTT
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 TTCCCTGTGGGATTTTTAAAACAGATATATTTTATTATTATTGTGACAAAATGTT

Qy 1369 AATGG 1373
              |||||
Db 1321 AATGG 1325

RESULT 5
LOCUS   AF055872   1368 bp   mRNA   linear   PRI 04
DEFINITION Homo sapiens Apo3/DR3 ligand (APO3L) mRNA, complete cds.
ACCESSION   AF055872
VERSION   AF055872.1 GI:3108230
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1368)
  AUTHORS   Marsters,S.A., Sheridan,J.P., Pitti,R.M., Brush,J., Goddar
            Ashkenazi,A.
  TITLE   Identification of a ligand for the death-domain-containing
            Apo3
  JOURNAL   Curr. Biol. 8 (9), 525-528 (1998)
  MEDLINE   98228355
   PUBMED   9560343
REFERENCE   2 (bases 1 to 1368)
  AUTHORS   Marsters,S.A., Sheridan,J.P., Pitti,R.M., Brush,J., Goddar
            Ashkenazi,A.
  TITLE   Direct Submission
  JOURNAL   Submitted (25-MAR-1998) Molecular Oncology, Genentech, 1 D
            South San Francisco, CA 94080, USA
FEATURES   Location/Qualifiers
     source   1..1368
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17p13"
     gene   1..1368
                     /gene="APO3L"
     CDS   58..807
                     /gene="APO3L"
                     /function="binds to the death-domain-containing r
                     Apo3/DR3"
                     /note="TNF homolog"
                     /codon_start=1
                     /product="Apo3/DR3 ligand"
                     /protein_id="AAC39724.1"
                     /db_xref="GI:3108231"
                     /translation="MAARSQRRRGRRGREPGTALLVPLALGLGLALACL
                     GSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQDPAPPLNRLVRPRR
                     ARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYNRQIG
                     YYLYCQVHFDEGKAVYLKLDLLVDGVLAIRCLEEFSATAASSLGPQLRL
                     RPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH"

ORIGIN

Query Match   96.2%; Score 1320.2; DB 9; Length 1368;
Best Local Similarity 99.8%; Pred. No. 8.4e-232;
Matches 1322; Conservative 0; Mismatches 3; Indels 0; Gaps 0; G

Qy   49 CGATCCCTCCGGTCCCGGATGGGGGGCGGTGGGTGAGGCGGTGAGGCCCCCGCGGCGCCCC
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CGATCCCTCCGGTCCCGGATGGGGGGCGGTGGGTGAGGCGGTGAGGCCCCCGCGGCGCCCC

Qy 109 GCCGGCCCCGTCCGAGCCCAGAGCGGCCAGAGGCGGAGGGGGCCCGGGGGCCACCCGCCCT
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GCCGGCCCCGTCCGAGCCCAGAGCGGCCAGAGGCGGAGGGGGCCCGGGGGCCACCCGCCCT

[alignment blocks continue]

             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                 GCTCGCGGCTGGGCCTGGGGCCTGCGCCTCCGGCTGGCCGTG 180

TTTGGGAGCGGGGACCGGATCGCTCGCGTCGTCCGCGGAGGCTGCCCAGGAGGCTGGTG 288
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCGAAGAGGCGGGATCGGCGGGCTGGGCTGCGCCATTAGAAGTCATCCACGACCTGACAGGACGGA 420
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCGAAGAGGCGGGATCGGCGGGCTGGGCTGCGCCATTAGAAGTCATCCACGACCTGACAGGACGGA 468

AGGCAGGTGTGGACGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCC 528
||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGGCAGGTGTGGACGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCC 480

TCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGCTGGGCTCTAC 588
||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTAC 540

GTACTGTCAGTGCACTTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGACTTG 648
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTACTGTCAGTGCACTTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGACTTG 600

GGATGGTGTGTGCTGGCCCTGCGCTGCCTGGAGGAATTCTCAGCCACTGCGCGCCAGT 708
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGATGGTGTGTGCTGGCCCTGCGCTGCCTGGAGGAATTCTCAGCCACTGCGCGCGAGT 660

CGGGGCCCCCAGCTCCGCCCTGTGCCAGGTGTCTGGCCGTGTTGGCCCTGGCCCTGGCCGTG 768
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CGGGGCCCCCAGCTCCGCCCTGCGCCAGGTGTCTGGCGTGTCTGGCCCTGTGGCCCTGGCCCAGGG 720

CCCTGCGGATCCGCACCCTCCCCGGGCCCATCTCAAGGCTGCCCCCTTCCTCACC 828
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCTGCGGATCCGCACCCTCCCCGGGCCCATCTCAAGGCTGCCCCCTTCCTCACC 780

CCGGACTCTTCCAGGTTCACTGAGGGGCCCTGGTCTCCCCACAGTCGTCCCAGGCT 888
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCGGACTCTTCCAGGTTCACTGAGGGGCCCTGGTCTCCCCAGTCGTCGTCCCAGGCT 840

GCTCCCCTGACAGAGTCTCTGGGCACCCGAGTCCCCCGTGCCCCAGCCGCGCT 948
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCTCCCCTGACAGAGTCTCTGGGCACCCGAGTCCCCCGTGCCCCACCCTCAGCCGCGCT 900

GCTCCAGACCTGCGCCCCTCCCTCTAGAGGCTGCCTGGCCGTTCACGTGTTTCCA 1008
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCTCCAGACCTGCGCCCCTCCCTCTAGAGGCTGCCTGGCCGTTCACGTGTTTCCA 960

ACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCCACCGCCCACTCTCCA 1068
||||||||||||||||||||||||||||||||||||||||||||||||||||||
ACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCCACCGCCCACTCTCCA 1020

ACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCCCCCTG 1128
||||||||||||||||||||||||||||||||||||||||||||||||||||||
ACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCCCCCTG 1080

CAGACCCCCAGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGTCCAGAAG 1188
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAGACCCCCAGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGTCCAGAAG 1140

CACTTCAGGCACTAAGAGGGGGCTGGACCTGGCCAGGAAGCCAAAGAGACTGGGC 1248
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CACTTCAGGCACTAAGAGGGGGCTGGACCTGGCCAGGAAGCCAAAGAGACTGGGC 1200

GCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAA 1308
||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db    1201 CTAGGCCAGGAGTTCCCAAATGTGAGGGGGCGAGAAACAAGACAAGCTCTCCCT
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1309 TTCCCTGTGGATTTTAAAACAGATATTATTTTATTATTGTGACAAAATGT
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 TTCCCTGTGGATTTTTAAAACAGATATTATTTTATTATTATTGTGACAAAATGT

Qy    1369 AATGG 1373
           |||||
Db    1321 AATGG 1325

RESULT 6
AF030099
LOCUS          AF030099                    1306 bp    mRNA        linear   PRI   2
DEFINITION     Homo sapiens TWEAK mRNA, complete cds.
ACCESSION      AF030099
VERSION        AF030099.1  GI:2707218
KEYWORDS       .
SOURCE         Homo sapiens (human)
  ORGANISM     Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutel
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
REFERENCE      1  (bases 1 to 1306)
  AUTHORS      Chicheportiche,Y., Bourdon,P.R., Xu,H., Hsu,Y.M., Scott,H
               Hession,C., Garcia,I. and Browning,J.L.
  TITLE        TWEAK, a new secreted ligand in the tumor necrosis factor
               that weakly induces apoptosis
  JOURNAL      J. Biol. Chem. 272 (51), 32401-32410 (1997)
  MEDLINE      98070415
  PUBMED       9405449
REFERENCE      2  (bases 1 to 1306)
  AUTHORS      Bourdon,P., Hession,C., Tizard,R. and Browning,J.
  TITLE        Direct Submission
  JOURNAL      Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridg
               Cambridge, MA 02142, USA
FEATURES             Location/Qualifiers
     source          1..1306
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17p13"
                     /tissue_type="tonsil"
     CDS             18..767
                     /note="ligand in the TNF family; secreted protei
                     codon not verified experimentally"
                     /codon_start=1
                     /product="TWEAK"
                     /protein_id="AAC51923.1"
                     /db_xref="GI:2707219"
                     /translation="MAARRSQRRRGRRGEPGTALLVPLALGLGLALAC
                     GSRASLSAQEPAQELVAEEDQDPSELNPQTEESQDPAFFLNRLVRPR
                     ARRAIAAHYEVHPRPGQDAQAGVDGTVSGWEEARINSSSPLRYNRQI
                     YYLYCVHFDEGKAVYLKLDLLVDGVLALRCLEEFSATAASSLGPQLR
                     RPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH"

ORIGIN
               Query Match          93.6%;  Score 1285;  DB 9;  Length 1306;
               Best Local Similarity  100.0%;  Pred. No. 2.4e-225;
               Matches 1285;  Conservative  0;  Mismatches  0;  Indels   0;

Qy      89 CACAGCCCCGCCCGCCCATGGCGCCGTCGCAGCGCCGCGGCCGGAGGCGGAGGGGCGCCG
           ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CACAGCCCCGCCCGCCCCATGGCGCCGTCGCAGCGCCGCGGCCGGAGGCGGAGGGGCGCCG
Qy     149 AGCCGGCACCGCACCGCGCCTGCTCCGCTCGCCTCGCCTGCCTCGGCGCTGGCGCGCTGGC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AGCCGGCACCGCACCGCGCCTGCTCCGCTCGCCTCGCCTGCCTCGGCGCTGGCGCGCTGGC
Qy     209 TCGGCCTCCTGCTGGCCGTGGCCGTTGGGGAGGCGGGCATCGCTGTCCGCCCA
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TCGGCCTCCTGCTGGCCGTGGCCGTTGGGGAGGCGGGCATCGCTGTCCGCCCA
```

Qy   1349 ATTGTGACAAAATGTTGATAAATGG 1373
           ||||||||||||||||||||||||
Db   1261 ATTGTGCAAAATGTTGATAAATGG 1285

RESULT 7
AR140407
LOCUS   AR140407   1236 bp   DNA   linear   PAT 1(
DEFINITION   Sequence 1 from patent US 6207642.
ACCESSION   AR140407
VERSION   AR140407.1   GI:14482903
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE   1 (bases 1 to 1236)
  AUTHORS   Wiley,S.R.
  TITLE   Member of the TNF family useful for treatment and diagnosi
             disease
  JOURNAL   Patent: US 6207642-A 1 27-MAR-2001;
FEATURES   Location/Qualifiers
     source 1..1236
             /organism="unknown"
             /mol_type="unassigned DNA"
ORIGIN

Query Match   89.3%; Score 1226.4; DB 6; Length 1236;
Best Local Similarity 99.5%; Pred. No. 1.3e-214;
Matches 1230; Conservative   0; Mismatches   6; Indels   0; C

Qy   106 ATGGCCGCCCGTCGGAGCCAGAGCCGGAGGCGGCACCGCCCCGGGGGAGCCCGGCACCGC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGGCCGCCCGTCGGAGCCAGAGCCGGAGGCGGCACCGCCCCGGGGGAGCCCGGCACCGC

Qy   166 CTGGTCCCGCTCGCGCTGGCGCTGGCGCTGGGCGCTGGCCTGGCCTCGGCCTCGGCCTCCTGCT
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTGGTCCCGCTCGCGCTGGCGCTGGGCGCTGGCCTGGCCTCGGCCTCGGCCTCCTGCT

Qy   226 GTGGTCAGTTTGGGAGCCGGCGGCATCGTGTCCGCCCAGGAGCCTGCCCAGGAGGA
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTGGTCAGTTTGGGGAGCCGGCGGCATCGTGTCCGCCCAGGAGCCTGCCCAGGAGGA

Qy   286 GTGGCAGAGGAGGAGACCAGGACCCGTCGGAACTGAATCCCAGACAGAAGAAAGCCA
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GTGGCAGAGGAGGAGACCAGGACCCGTCGGAACTGAATCCCAGACAGAAGAAAGCCA

Qy   346 CCTGCGGCCTTTCTGAACCGACTAGTTCGGCCTCGAGAAGTGCACCTAAAGCCCC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CCTGCGGCCTTTCTGAACCGACTAGTTCGGCCTCGAGAAGTGCACCTAAAGCCCG

Qy   406 ACACGGCTCGAAGAGCAGGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACA
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ACACGGCTCGAAGAGCAGGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACA

Qy   466 GGAGCGCAGGCAGTGTGGACGGGACAGTGAGTGAGTGGCTGGGAGGAAGCCAGAATCAA
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GGAGCGCAGGCAGTGTGGACGGGACAGTGAGTGAGTGGCTGGGAGGAAGCCAGAATCAA

Qy   526 TCCAGCCCTGCGCTACAACCGCCAGATCGGGAGTTTATAGTCACCGGGCTGG
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TCCAGCCCTGCGCTACAACCGCCAGATCGGGAGTTTATAGTCACCGGGCTGG

Qy   586 TACTACCTGTACTGTCAGGTGCACTTTGATGAGGGAGGAAGGCGTGTCACCTGAAGCT
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TACTACCTGTACTGTCAGGTGCACTTTGATGAGGGAGGAAGGCGTGTCACCTGAAGCT

Qy   706 AGTTCCCTCGCGCCCGCCCGCTCCGCCCTCGCCAGGTGTCTGGCCCTGCGCG
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 AGTTCCCTCGCGCCCGCCCGCTCCGCCCTCGCCAGGTGTCTGGCCCTGCGCG

CCAGGAGGAGCTGGTGGCAGAGGAGGAGGACCAGGACCCGTCGGAACTGAATCCCAGA 328
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCAGGAGGAGCTGGTGGCAGAGGAGGAGGACCAGGACCCGTCGGAACTGAATCCCAGA 240

AGAAAGCCAGGATCCTGCGCCTTTCTGAACCGACTAGTTCGGCCTCGAGAAGTG 388
|||||||||||||||||||||||||||||||||||||||||||||||||||||
AGAAAGCCAGGATCCTGCGCCTTTCTGAACCGACTAGTTCGGCCTCGCAGAGTG 300

TAAAGGCCGGAAAACACGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATC 448
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
TAAAGGCCGGAAAACACGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATC 360

ACCTGGACAGGACGGAGCGCAGGCAGGTGTGGACAGTGAGTGGCTGGGAGG 508
||||||||||||||||||||||||||||||||||||||||||||||||||
ACCTGGACAGGACGGAGCGCAGGCAGGTGTGGACAGTGAGTGGCTGGGAGG 420

CAGAATCACACAGTCCAGCCTCTGCGCTACAACCGCCAGATCGGGAGTTTATAG 568
|||||||||||||||||||||||||||||||||||||||||||||||||||||
CAGAATCAACAGTCCAGCCTCTGCGCTACAACCGCCAGATCGGGAGTTTATAG 480

CCGGGCTGGGCCTCTACTACCTGTACTGTCAGTGCACTTTGATGAGGGAAGGCTG 628
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCGGGCTGGGCCTCTACTACCTGTACTGTCAGGTGCACTTTGATGAGGGAAGGCTG 540

CCTGAAGCTGGACTTGCTGGTGGATGGTGTGCTGGCCCTGCCCTGCCTGGAGCGAAT 688
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCTGAAGCTGGACTTGCTGGTGGATGGTGTGCTGGCCCTGCCCTGCCTGGAGCGAAT 600

AGCCACTGCGGCCAGTTCCCTCGCGCCCCAGCTCCGCCCTCTGCCAGGTGTCTGGGC 748
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGCCACTGCGGCCAGTTCCCTCGCGCCCCAGTTCCCTCTGCCCAGTGTCTGGGC 660

GGCCCTGCGCGCCAGGGTCCTCCCTGCGATCGCAGCACCCTCCCCTGGGCCCATCTCA 808
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGCCCTGCGCGCCAGGGTCCTCCCTGCGATCGCAGCACCCTCCCCTGGGCCCATCTCA 720

TGCCCCCTTCCTCACCTACTTCGGACTCTTCCAGGTTCACTGAGGGCCCTGGTCT 868
|||||||||||||||||||||||||||||||||||||||||||||||||||||
TGCCCCCTTCCTCACCTACTTCGGACTCTTCCAGGTTCACTGAGGGCCCTGGTCT 780

ACAGTCGTCCCAGGCTGCGCGCTCCCCCTGGCAGCTCTCTGGGCACCCGGTCCCCT 928
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
ACAGTCGTCCCAGGCTGCGCGCTCCCCCTGGCAGCTCTCTGGGCACCCGGTCCCCT 840

CCCACCCCCAGCCGCTCTTGCTCCAGACCTGCCCCTCCCCTCTAGAGGCTGCCTGG 988
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCACCCCCAGCCGCTCTTGCTCCAGACCTGCCCCTCCCCTCTAGAGGCTGCCTGG 900

GTTCACGTGTGTTTCCATCCCACATAAATACAGTATTCCCACTCTTATTCTTACAAACT 1048
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTTCACGTGTGTTTCCATCCCACATAAATACAGTATTCCCACTCTTATTCTTACAAACT 960

CACCGCGCCACTCTCACCACTCTACTAGCTCCCCAATCCCTGACCCTTTGAGGCGCCCCA 1108
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CACCGCCCACTCTCACCACTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCGCCCCA 1020

TCTCGACTCCCCCCTGGCCACAGACCCCAGGCATTGTTCACTGTACTCTGTG 1168
|||||||||||||||||||||||||||||||||||||||||||||||||||
TCTCGACTCCCCCCTGGCCACAGACCCCAGGCATTGTTCACTGTACTCTGTG 1080

AGGATGGGTCAGAAGACCCACTCAGGCACTAAGAGGGGCTGGACCTGGCGGGCA 1228
|||||||||||||||||||||||||||||||||||||||||||||||||||||
AGGATGGGTCAGAAGAACCCACTCAGGCACTAAGAGGGGCTGGACCTGGCGGGCA 1140

GCCAAAGAGATGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAG 1288
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAG 1200

GCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTATTTTTATTATT 1348
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTATTTTTATTATT 1260

/db_xref="taxon:9606"

ORIGIN

Query Match   89.3%;   Score 1226.4; DB 6;   Length 1236;
Best Local Similarity 99.5%;   Pred. No. 1.3e-214;
Matches 1230; Conservative   0; Mismatches   6; Indels   0;

```
Qy    106 ATGGCCGCCCGTCGGAGCCAGAGGCCGGAGGGGGCGCCGGGGGAGCCGGGCACCG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGCCGCCCGTCGGAGCCAGAGGCCGGAGGGGGCGCCGGGGGAGCCGGGCACCG

Qy    166 CTGGTCCCGTCGGCGTCGGCCTGGGCCTGGCCTGGCCGCGTCGGCCTCCGCCTCCTGC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTGGTCCCGTCGGCGTCGGCCTGGGCCTGGCCTGGCCGCGTCGGCCTCCGCCTCCTGC

Qy    226 GTGGTCAGTTGGGAGAGCCGGGATCGCTGTCGCCAGGAGCTGCCCAGGAGG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTGGTCAGTTGGGAGAGCCGGGATCGCTGTCGCCAGGAGCTGCCCAGGAGG

Qy    286 GTGGCAGAGGAGGACCAGGACCCGTCGGAACTGAATCCCAGACAGAAGAAGAAGCC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTGGCAGAGGAGGACCAGGACCCGTCGGAACTGAATCCCAGACAGAAGAAGAAGCC

Qy    346 CCTGCGCCTTCCTGAACCGACTAGTTCGGCCTGCAGAGTGCACCTAAAGGCC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCTGCGCCTTCCTGAACCGACTAGTTCGGCCTGCAGAGTGCACCTAAAGGCC

Qy    406 ACACGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGAC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ACACGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGAC

Qy    466 GGAGCGCAGGCAGTGTGACGGGACAGTGAGTGCTGGGAGGAAGCCAGAATCA
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GGAGCGCAGGCAGTGTGACGGGACAGTGAGTGCTGGGAGGAAGCCAGAATCA

Qy    526 TCCAGCCCTCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCCAGCCCTCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTG

Qy    586 TACTACCTGTACTGTCAGGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TACTACCTGTACTGTCAGGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGC

Qy    646 TTGCTGGTGATGGTGTGTCCGCTGCCCGCTGCCCTGCCCTCTGGAGAATTCTCAGCCACTG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TTGCTGGTGATGGTGTGTCCGCTGCCCGCTGCCCTGCCCTCTGGAGAATTCTCAGCCACTG

Qy    706 AGTTCCTCTGGGCCCCAGCTCTCCGCCTCTGCGACAGTCGTCTGGGCCCTGC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AGTTCCTCTGGGCCCCAGCTCTCCGCCTCTGCGACAGTCGTCTGGGCCCTGC

Qy    766 GGGTCCTCCCGCGATCCGACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCT
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GGGTCCTCCCGCGATCCGACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCT

Qy    826 ACCTACTTCGGACTCTTCCAGGTTCACTGAGGGGCCCTGCTCTCCCCACAGTCGT
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 ACCTACTTCGGACTCTTCCAGGTTCACTGAGGGGCCCTGCTCTCCCCGCAGTCGT

Qy    886 GCTGCCGGCTCCCCTCGACAGCTCTCGGCCACCCGGTCCCCTCGCCCACCCT
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GCTGCCGGCTCCCCTCGACAGCTCTCGGCCACCCGGTCCCCTCGCCCACCCT

Qy    946 GCTCTTTGCTCCAGACCTGCCCCTTCCCCTAGAGGCTGCTGGGCCTGTTCACGT
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 GCTCTTTGCTCCAGACCTGCCCCTTCCCCTTAGAGCTGCTGGGCCTGTTCACGT

Qy   1006 CCATCCCACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCCCC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCATCCCACATAAATACAGTATTCCCACTCTTATCTTACAACACCCCACCGCCCCC

Qy   1066 CCACCTACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCAGTGATCTCGACT
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
```

(Right side column)

```
Qy        CCCTCCCTGCGGGATCCGCACCCTCCCCTGGGCGCCATCTCAAGGCTGCCCCTTCCTC 825
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        CCCTCCCTGCGGGATCCGCACCCTCCCCTGGGCGCCATCTCAAGGCTGCCCCTTCCTC 720

Qy        TACTTCGGACTCTTCCAGGTTCACTGAGGGGCCCTGGTCTCCCCAGTCGTCCCAG 885
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        TACTTCGGACTCTTCCAGGTTCACTGAGGGGCCCTGGTCTCCCCAGTCGTCCCAG 780

Qy        GCCGGCTCCCCTCGACAGCTCTCGGCCACCCGGTCCCCTCTGCCCACCCTCAGCC 945
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        GCCGGCTCCCCTCGACAGCTCTCGGCCACCCGGTCCCCTCTGCCCACCCTCAGCC 840

Qy        CTTTGCTCCAGACCTGCCCCTTCCCTTAGAGGCTGCCTGGCCTGTTCACGTGTTTT 1005
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        CTTTGCTCCAGACCTGCCCCTTCCCTTAGAGGCTGCCTGGCCTGTTCACGTGTTTT 900

Qy        CCCACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCCACTCT 1065
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        CCCACATAAATACAGTATTCCCACTCTTATCTTACAACACCCCACCGCCCACTCT 960

Qy        CCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCAGTGATCTCGACTCCCCC 1125
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        CCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCAGTGATCTCGACTCCCCC 1020

Qy        GCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGTCCAG 1185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        GCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGTCCAG 1080

Qy        ACCCCACTTCAGGCACTAAGAGAGGGGCTGGACCTGGGCGCAGGAGGCCAAAGAGACTG 1245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        ACCCCACTTCAGGCACTAAGAGAGGGGCTGGACCTGGGCGCAGGAGGCCAAAGAGACTG 1140

Qy        CTAGGCCAGGAGTTCCCAAATGTGAGGGGCCGAGAAACAAGACAAGCTCCTCCCTTGA 1305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        CTAGGCCAGGAGTTCCCAAATGTGAGGGGCCGAGAAACAAGACAAGCTCCTCCCTTGA 1200

Qy        TTCCCTGTGTGGATTTTTAAAACAGATATTATTTT 1341
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        TTCCCTGTGTGGATTTTTAAAACAGATATTATTTT 1236
```

124
c of the tnf family useful for treatment and diagnosis of
e.
124
124.1 GI:22602730
01513626-A/1.
sapiens (human)
sapiens
yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
lia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ases 1 to 1236)
,S.R.
r of the tnf family useful for treatment and diagnosis of
:: JP 2001513626-A 1 04-SEP-2001;
r LABORATORIES
JP 2001513626-A/1
04-SEP-2001
12-FEB-1998 JP 1998535077
12-FEB-1997 US 08/798692,10-FEB-1998 US   09/021706 PI
N R WILEY
C12Q1/68,C12N15/63,C07K14/525,C07K16/28,G01N33/50,G01N33/574,
A61K38/19
Strandedness: Single;
Topology: Linear;
Key   Location/Qualifiers
1..1236
      /organism="Homo sapiens"
      /mol_type="genomic DNA"

1236 bp   DNA   linear   PAT 27-AUG-2002

Web site: http://www.nisc.nih.gov/
Contact: nisc mgc@nhgri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Broc
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legas
Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McClos
McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touch
Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggin
Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can
through the I.M.A.G.E. Consortium/LLNL at: http://image.il
Series: IRAL Plate: 30 Row: p Column: 5
This clone was selected for full length sequencing because
passed the following selection criteria: matched mRNA gi:

FEATURES   Location/Qualifiers
     source   1..1642
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:20669 IMAGE:4766071"
                     /tissue_type="Primary B-Cells from Tonsils"
                     /clone_lib="NIH_MGC_48"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     gene   1..1642
                     /gene="TNFSF12"
                     /note="synonyms: TWEAK, DR3LG, APO3L, MGC20669,
                     /db_xref="LocusID:8742"
                     /db_xref="MIM:602695"
     CDS   97..501
                     /codon_start=1
                     /product="tumor necrosis factor (ligand) superfami
                     member 12, isoform 2"
                     /protein_id="AAH19047.1"
                     /db_xref="GI:17512139"
                     /db_xref="LocusID:8742"
                     /translation="MAARRSORRRGRRGEPGTALLVPLALGLGLALACLG
                     GSRASLSAQEPAQEELVAEEDQPSELNPQTEESQDPAPFLNRLVRPRRS
                     ARRAIAAHYEVHPRPGQDGAQADGGYTTCLRP"

ORIGIN
     Query Match   79.9%;   Score 1096.8;   DB 9;   Length 1642;
     Best Local Similarity   84.8%;   Pred. No. 6.9e-191;
     Matches 1347;   Conservative   0;   Mismatches   2;   Indels 239;   Ga

Qy   25 TCCGCCCGCCGGCTCCCCCATCCCTGGGTCCCGGATGGGGGGCGGTG
            |||||||||||||||||||||||||||||||||||||||||||||||
Db   16 TCCGCCCGCCGGCTCCCCCATCCCTGGGTCCCGGATGGGGGGCGGTG.

Qy   85 CAGGCACAGCGCCCCGCGCCCCCATGGCGCCCGTCGGAGCCAGAGGCCGGAGGGGGCG(
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   76 CAGGCACAGCGCCCCGCGCCCCCATGGCGCCCGTCGGAGCCAGAGGCCGGAGGGGGCG(

Qy   145 GGGGAGCCGGGCACCGCCCCTGCTGGTCCCGGCTGGGGGCCTGGGCTG(
            |||||||||||||||||||||||||||||||||||||||||||||||
Db   136 GGGGAGCCGGGCACCGCCCCTGCTGGTCCCGGCTGGGGGCCTGGGCTG(

Qy   205 TGCCTGGCCTCCTGCTGCCGTGTCAGTTTGGGGAGCCGGGAGCATCGCGTGTCCGCC(
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   196 TGCCTGGCCTCCTGCTGCCGTGTCAGTTTGGGGAGCCGGGAGCATCGCGTGTCCGCC(

Qy   265 GAGGCTGCCCAGGAGGAGCTGGTGGCAGAGGAGGACCAGGACCCGTCGGAACTGAAT(
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   256 GAGGCTGCCCAGGAGGAGCTGGTGGCAGAGGAGGACCAGGACCCGTCGGAACTGAAT(

Qy   325 CAGACAGAAGAAAGCCAGATCCTGCGCCTTCTGAACCGACTAGTTCGGCCTCGCA
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   316 CAGACAGAAGAAAGCCAGATCCTGCGCCTTCTGAACCGACTAGTTCGGCCTCGCA

Qy   385 AGTGCACCTAAAGCCCGGAAAACACGGCGTCGAAGAGCGATCGCAGCCCATTATGAAG
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

CTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCCCCC 1020
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     CCACAGACCCCCAGGGCATTGTTCACTGTACTCTGTGGGCAAGGATGGGGTCCAG 1185
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     CCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAG 1080

     CCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCCGGCGGCAGCAAAGCCAAAAGAGACTG 1245
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     CCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCCGGCGGCAGGAGGCCAAAGAGACTG 1140

     CAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGA 1305
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     CAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGA 1200

     CCCTGTGTGGATTTTTTAAAACAGATATTATTTT 1341
     ||||||||||||||||||||||||||||||||||
     CCCTGTGTGGATTTTTAAAACAGATATTATTTT 1236

     7   1642 bp   mRNA   linear   PRI 04-OCT-2003
piens tumor necrosis factor (ligand) superfamily, member 12,
ipt variant 2, mRNA (cDNA clone MGC:20669 IMAGE:4766071),
e cds.
7
7.2   GI:33874709

piens (human)
piens
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 1642)
erg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
1,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
nko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
on,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
i,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
a,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
i,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
i,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
.A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
i,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
eld,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
,A., Schein,J.E., Jones,S.J. and Marra,M.A.
on and initial analysis of more than 15,000 full-length
d mouse cDNA sequences
tl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

es 1 to 1642)
rg,R.
ubmission
d (07-DEC-2001) National Institutes of Health, Mammalian
lection (MGC), Cancer Genomics Office, National Cancer
e, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

Project URL: http://mgc.nci.nih.gov
9, 2003 this sequence version replaced gi:17512138.
   MGC help desk
gapbs-r@mail.nih.gov
rocurement: Louis Staudt
rary Preparation: Rubin Laboratory
rary Arrayed by: The I.M.A.G.E. Consortium (LLNL)
encing by: National Institutes of Health Intramural
ng Center (NISC),
burg, Maryland;

us-09-245-198a-3.rge

```
Qy    1286 AAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATATTT
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1516 AAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATATTT

Qy    1346 ATTATTGTGACAAAATGTTGATAAATGG 1373
           ||||||||||||||||||||||||||||
Db    1576 ATTATTGTGACAAAATGTTGATAAATGG 1603
```

RESULT 10
AC016876/c
LOCUS   AC016876   60268 bp   DNA   linear   PRI 3
DEFINITION   Homo sapiens chromosome 17, clone RP11-186B7, complete se
ACCESSION   AC016876
VERSION   AC016876.10 GI:24431829
KEYWORDS   HTG.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutel
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
REFERENCE   1 (bases 1 to 60268)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE   Homo sapiens chromosome 17, clone RP11-186B7
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 60268)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., And
             Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhga
             Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A
             Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L.,
             Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
             Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Ho
             Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., K
             Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
             McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,
             Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,
             Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Seve
             Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,
             Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., W
             Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (08-DEC-1999) Whitehead Institute/MIT Center for
             Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3 (bases 1 to 60268)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anders
             Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalte
             Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A
             Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dod
             Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
             Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
             Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat
             Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,
             Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
             McCarthy,M., Meldrim,J., Meneus,L., Nicol,R., Norbu,C., Norma
             Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norma
             O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,
             Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Ro
             Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Sev
             Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Ta
             Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,
             Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,C., Zainc
             Zembek,L., Zimmer,A. and Zody,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (08-OCT-2002) Whitehead Institute/MIT Center for
             Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   4 (bases 1 to 60268)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderso
             Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalte
             Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A
             Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodc
             Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
             Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
             Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,
```

```
      GCACCTAAAGGCCGAGAAAACACGGGCTCGAAGAGCGATCGCGAGCCCATTATGAAGTT 435

      CCACGACCTGGACAGGACGGAGCGCAGG------------------------------ 475
      ||||||||||||||||||||||
      CCACGACCTGGACAGGACGGAGCGCAGGCGCAGGCGAGTGAGCGTTACACAACTTGTCTGAGG 495

      TGAGATACTAAGTGGTGGAGCCAAGAGATTTGAACCCAGCTAGAAATGTGCCTGTGTACT 555
      --------------------------------------------------------- 475

      CATGGTTCTGCATGAAAGGCGAGGTGCTGGTGCAGGGGTGAGGGGTCCATGCAGGGG 615
      --------------------------------------------------------- 475

      CATCCAAAAAGGGGAGAGGGAAGTTTCCAGAAAGAAGAGAGGACACATCTCCCACCA 675
      -----------------CAGGTGTGACGGGACAGTGAGTGGCTGGG 505
                        ||||||||||||||||||||||
      CCAGAGGGGTCAAAGGAGAACAAGAGACCAGTGTGACGGGACAGTGAGTGGCTGGG 735

      AGCCAGAATCAACAGCTCCAGCCCTCTGCCTACAACCGCCAGAGTCGGGGAGTTTA 565
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
      AGCCAGAATCAACAGCTCCAGCCCTCTGCCTACAACCGCCAGAGTCGGGGAGTTTA 795

      CACCCGGGCTGGGCTCTACTACTTGTCAGGTGCACTTTGATGAGGGGAAGG 625
      ||||||||||||||||||||||||||||||||||||||||||||||||||
      CACCCGGGCTGGGCTCTACTACTTGTCAGGTGCACTTTGATGAGGGGAAGG 855

      CTACTTGAAGCTGGACTTGCTGGTGGATGGTGTGCTGGCCCTGCCCTGCCCTGGAGG 685
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
      CTACTTGAAGCTGGACTTGCTGGTGGATGGTGTGCTGGCCCTGCCCTGCCCTGGAGG 915

      CTCAGCCACTGCGGCCAGTTCCCTCGGGGCCCCAGCCTTCCGGCCCTTGCCAGGTGTCTG 745
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
      CTCAGCCACTGCGGCCAGTTCCCTCGGGGCCCCAGCCTTCCGGCCCTTGCCAGGTGTCTG 975

      GTTGGGCCCTGCGGCCAGGGTCCTCCCTGCGGATCCGCACCCTCCCCTGGGCCCATC 805
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
      GTTGGGCCCTGCGGCCAGGGTCCTCCCTGCGGATCCGCACCCTCCCCTGGGCCCATC 1035

      GGCTGCCCCCTTCCTCACCTACTTCGGACTCTTCCAGGTTCACTGAGGGGCCCTGG 865
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
      GGCTGCCCCCTTCCTCACCTACTTCGGACTCTTCCAGGTTCACTGAGGGGCCCTGG 1095

      CCCACAGTCGTCCCAGGCTGCCGGCTCCCCTCCGGCAGCTCTTCTGGGCCACCGGGTCC 925
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
      CCGCAGTCGTCCCAGGCTGCCGGCTCCCCTCCGGCAGCTCTTCTGGGCCACCGGGTCC 1155

      TGCCCCACCCTCAGCCGCTCTTTGCTCCAGACCTGCCCCCTCCCCTCTAGAGGCTGCC 985
      |||||||||||||||||||||||||||||||||||||||||||||||||||||
      TGCCCCACCCTCAGCCGCTCTTTGCTCCAGACCTGCCCCCTCCCCTCTAGAGGCTGCC 1215

      CTGTGTTCACGTGTTTTCCATCCCACATAAATACAGTATTCCCACTCTTATCTTACA 1045
      |||||||||||||||||||||||||||||||||||||||||||||||||||||
      CTGTGTTCACGTGTTTTCCATCCCACATAAATACAGTATTCCCACTCTTATCTTACA 1275

      CCCACACGCCCACTCTCCACCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGGCCC 1105
      |||||||||||||||||||||||||||||||||||||||||||||||||||||
      CCCACCCGCCCACTCTCCACCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGGCCC 1335

      GGATCTTCGACTCCCCCCTGGCCACAGAGACCCCACTTCAGGGCATTGTGTTCACTGTACTCT 1165
      |||||||||||||||||||||||||||||||||||||||||||||||||||||
      GGATCTTCGACTCCCCCCTGGCCACAGAGACCCCACTTCAGGGCATTGTGTTCACTGTACTCT 1395

      CCAAGGATGGTCCAAGAAGCACCCCAGTTCAGGCACTAAGGACTCGGACCCTGCGCG 1225
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
      CCAAGGATGGTCCAGAGAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCCTGCGCG 1455

      AAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAAC 1285
      ||||||||||||||||||||||||||||||||||||||||||||||||||
      AAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAAC 1515
```

A., Karatas,A., Kells,C., Landers,T., Levine,R.,
ad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
ws,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
on,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
, Rogov,P., Ronan,J., Roy,A., Schauer,S., Schupback,R.,
,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
ovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
s,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Submission
ted (31-OCT-2002) Whitehead Institute/MIT Center for Genome
ch, 320 Charles Street, Cambridge, MA 02141, USA
31, 2002 this sequence version replaced gi:23592141.
peats were identified using RepeatMasker.
A.F.A. & Green, P. (1996-1997)
/ftp.genome.washington.edu/RM/RepeatMasker.html
------- Genome Center
nter: Whitehead Institute/ MIT Center for Genome Research
nter code: WIBR
o site: http://www-seq.wi.mit.edu
ntact: sequence_submissions@genome.wi.mit.edu
------- Project Information
iter project name: L3849
iter clone name: 186_B_7
----

e first 60.3 kilobases of this clone are being submitted.
ainder overlaps accession number AC113189 [WICGR project

Location/Qualifiers
1..60268
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-186B7"
/clone_lib="RPCI-11 Human Male BAC"
complement(7920..8209)
/rpt_family="AluSq"
complement(8822..9119)
/rpt_family="AluSx"
complement(9356..9663)
/rpt_family="AluSq"
complement(12421..12733)
/rpt_family="AluSx"
14360..14438
/rpt_family="(TGG)n"
complement(17883..18086)
/rpt_family="MIR"
complement(18188..18483)
/rpt_family="AluSq"
22454..22535
/rpt_family="L2"
24275..24530
/rpt_family="MIR"
complement(24517..24655)
/rpt_family="MIR"
24839..24954
/rpt_family="L2"
complement(24973..25268)
/rpt_family="AluY"
complement(25273..25768)
/rpt_family="MER21B"
complement(25769..26057)
/rpt_family="AluSg"
complement(26058..26334)
/rpt_family="MER21B"
26344..26838
/rpt_family="L2"
complement(26839..27139)

repeat_region	/rpt_family="AluSq" 27140..27185
	/rpt_family="L2" 27405..27430
unsure	/note="<30 qual SNGL region" 27529..27550
unsure	/note="single clone coverage" complement(27552..27689)
repeat_region	/rpt_family="FLAM_C" 27568..27572
unsure	/note="<30 qual SNGL region" complement(27690..28000)
repeat_region	/rpt_family="AluSx" complement(28005..28312)
repeat_region	/rpt_family="AluSp" 30022..30088
repeat_region	/rpt_family="L2" 30817..30846
repeat_region	/rpt_family="GC_rich" 31023..31072
repeat_region	/rpt_family="(TCC)n" 32341..32378
repeat_region	/rpt_family="(CCG)n" 32571..32601
repeat_region	/rpt_family="GC_rich" 37977..38092
repeat_region	/rpt_family="L1MB5" 38094..38400
repeat_region	/rpt_family="AluSx" 38910..39189
repeat_region	/rpt_family="L2" complement(39218..39381)
repeat_region	/rpt_family="AluSg/x" 39394..39457
repeat_region	/rpt_family="L2" 39458..39748
unsure	/rpt_family="AluY" 39593..39597
	/note="<30 qual SNGL region" 39749..39759
repeat_region	/rpt_family="L2" 39760..40068
repeat_region	/rpt_family="AluY" 40069..40152
repeat_region	/rpt_family="L2" 40350..40630
repeat_region	/rpt_family="L2" 40718..40905
repeat_region	/rpt_family="MIR" 40926..41025
repeat_region	/rpt_family="L2" 41026..41341
repeat_region	/rpt_family="AluSp" 41342..41745
repeat_region	/rpt_family="L2" 41746..42046
repeat_region	/rpt_family="AluSx" 42047..42288
repeat_region	/rpt_family="L2" 42289..42581
repeat_region	/rpt_family="AluSx" 42582..42730
repeat_region	/rpt_family="L2" 43730..43859
repeat_region	/rpt_family="MER94" 44018..44132

Query Match   56.0%;   Score 768.2;   DB 9;   Length 60268;
Best Local Similarity 99.6%;   Pred. No. 8.1e-131;
Matches 770;   Conservative   0;   Mismatches   3;   Indels   0;   Ga

Qy   601 CAGGTGCACTTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGGAT
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Query Match   1.9%; Score 26; DB 6; Length 487;
Best Local Similarity 100.0%; Pred. No. 0.022; Indels 0;
Matches 26; Conservative 0; Mismatches 0;

Qy   1 ATGTCATTGTTAGACTTGAAATTTC 26
       |||||||||||||||||||||||||||
Db 427 ATGTCATTGTTAGACTTGAAATTTC 452

RESULT 52
BC060315
LOCUS   BC060315   495 bp   mRNA   linear   PRI 2
DEFINITION Homo sapiens cDNA clone IMAGE:3879577, partial cds.
ACCESSION   BC060315
VERSION   BC060315.1 GI:37805340
KEYWORDS
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
REFERENCE   1 (bases 1 to 495)
AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Sc
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bl
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Ho
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hon
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Pete
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., M
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE   Generation and initial analysis of more than 15,000 full
            human and mouse cDNA sequences
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002
MEDLINE   22382257
PUBMED   12477932
REFERENCE   2 (bases 1 to 495)
AUTHORS   Strausberg,R.
TITLE   Direct Submission
JOURNAL   Submitted (20-OCT-2003) National Institutes of Health, M
            Gene Collection (MGC), Cancer Genomics Office, National
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNI
            DNA Sequencing by: National Institutes of Health Intramu
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 208
            USA
REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL
            DNA Sequencing by: National Institutes of Health Intram
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., B
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Le
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McC
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., To
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wig
            Young,A., Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information
            through the I.M.A.G.E. Consortium/LLNL at: http://image

---

095   486 bp   DNA   linear   PAT 17-JUL-2003
nce characteristic to gene transcription controlled by
ia.
095
095.1 GI:33033865
2525081-A/10.
sapiens (human)
sapiens
yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
lia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ases 1 to 486)
nce characteristic to gene transcription controlled by hypoxia
t: JP 2002525081-A 10 13-AUG-2002;
 BIOTECH INC
Homo sapiens (human)
JP 2002525081-A/10
13-AUG-2002
27-AUG-1999 JP 2000571058
27-AUG-1998 US   60/098158,05-MAY-1999 US   60/132684 PI
INAT,RAMI SKALITER,ELENA FEINSTEIN
C12N15/09,A61K31/711,A61K45/00,A61K48/00,A61P9/10,A61P9/00,A61P9/10, PC
P43/00,
C12Q1/68,G01N33/53,G01N33/566,G01N33/68,G01N37/00,
C12N15/00
Sequence characteristic to gene transcription controlled by
            hypoxia
Key   Location/Qualifiers
source   1..486
             /organism='Homo sapiens (human)'.
      Location/Qualifiers
      1..486
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"

ilarity 100.0%; Score 26; DB 6; Length 486;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GTCATTGTTAGACTTGAAATTTC 26
|||||||||||||||||||||||
GTCATTGTTAGACTTGAAATTTC 425

 sapiens (human)

8082   487 bp   DNA   linear   PAT 14-JUN-2002
ence 729 from Patent WO0229103.
8082
8082.1 GI:21440787

 sapiens (human)
uryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
nalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
     Location/Qualifiers
     1..487
     /organism="Homo sapiens"
     /mol_type="unassigned DNA"
     /db_xref="taxon:9606"
     /note="EMBL/GenBank Accession No. AA306121"

res,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
s expression profiles in liver cancer
ent: WO 0229103-A 729 11-APR-2002;
E LOGIC INC (US)

```
PF   07-AUG-2000 JP 2000280989
PR   05-AUG-1999 US   60/147499
PI   JEAN BAPUTIST DUMAS MILNE EDWARDS,SEVELIN JOBERT,JEAN
     GIORDANO
PC   C12N15/09,C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N
     C12N1/21,
PC   C12N5/10,C12P21/02,C12P21/08,C12Q1/68,C12N15/00,C12N5
     C12N15/00
CC   EST and encoded human protein
FH   Key             Location/Qualifiers
FT   CDS             338..499.

FEATURES             Location/Qualifiers
     source          1..499
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

ORIGIN
Query Match      1.9%;   Score 26;  DB 6;   Length 499;
Best Local Similarity  100.0%;  Pred. No. 0.022;
Matches  26;  Conservative  0;  Mismatches  0;  Indels  0;  G

Qy    1 ATGTCATTGTTAGACTTTGAAATTTC 26
        ||||||||||||||||||||||||||
Db  473 ATGTCATTGTTAGACTTTGAAATTTC 498


RESULT 55
AX885407
LOCUS       AX885407    516 bp    DNA    linear    PAT 18
DEFINITION  Sequence 1270 from Patent EP1033401.
ACCESSION   AX885407
VERSION     AX885407.1  GI:40041537
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
  TITLE     Expressed sequence tags and encoded human proteins
  JOURNAL   Patent: EP 103401-A 1270 06-SEP-2000;
            Genset (FR)
FEATURES             Location/Qualifiers
     source          1..516
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             343..>516
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAE99545.1"
                     /db_xref="GI:40041538"
                     /translation="MEYERSPLERKGQTLCFHESEDLAEPVPQGYCIHS
                     RMSLLDFEISNCLP"

ORIGIN
Query Match      1.9%;   Score 26;  DB 6;   Length 516;
Best Local Similarity  100.0%;  Pred. No. 0.022;
Matches  26;  Conservative  0;  Mismatches  0;  Indels  0;  G

Qy    1 ATGTCATTGTTAGACTTTGAAATTTC 26
        ||||||||||||||||||||||||||
Db  478 ATGTCATTGTTAGACTTTGAAATTTC 503


RESULT 56
BD025017
LOCUS       BD025017    516 bp    DNA    linear    PAT 27
DEFINITION  Sequence tag and encoded human protein.
ACCESSION   BD025017
VERSION     BD025017.1  GI:22566240
KEYWORDS    JP 2001269182-A/1263.
```

```
IRAK Plate: 134 Row: c Column: 18
one was selected for full length sequencing because it
the following selection criteria: Hexamer frequency ORF
s.

Location/Qualifiers
1..495
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3879577"
/tissue type="Lung, large cell carcinoma"
/clone_lib="NIH MGC_68"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

arity   1.9%;   Score 26;  DB 9;   Length 495;
onservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

CATTGTTAGACTTTGAAATTTC 26
|||||||||||||||||||||
CATTGTTAGACTTTGAAATTTC 424


25      499 bp    DNA    linear    PAT 18-DEC-2003
ce 1162 from patent US 6639063.
25.1  GI:40168635

.
ified.
ses 1 to 499)
,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
and encoded human proteins
: US 6639063-A 1162 28-OCT-2003;
Location/Qualifiers
1..499
/organism="unknown"
/mol_type="genomic DNA"

arity   1.9%;   Score 26;  DB 6;   Length 499;
onservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

CATTGTTAGACTTTGAAATTTC 26
|||||||||||||||||||||
CATTGTTAGACTTTGAAATTTC 498


78      499 bp    DNA    linear    PAT 18-SEP-2002
d encoded human protein.
78.1  GI:23203896
2010789-A/1155.
apiens (human)
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 499)
,J.B.D.M., Jobert,S. and Giordano,J.E.
d encoded human protein
: JP 2002010789-A 1155 15-JAN-2002;
CORP
omo sapiens (human)
P 200201078 9-A/1155
5-JAN-2002
```

VERSION   BD111487.1 GI:23206305
KEYWORDS   JP 2002010789-A/3564.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutel
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
            1 (bases 1 to 522)
REFERENCE   Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
AUTHORS   EST and encoded human protein
TITLE   Patent: JP 2002010789-A 3564 15-JAN-2002;
JOURNAL   GENSET CORP
COMMENT   OS Homo sapiens (human)
            PN JP 2002010789-A/3564
            PD 15-JAN-2002
            PF 07-AUG-2000 JP 2000280989
            PR 05-AUG-1999 US   60/147499
            PI JEAN BAPUTIST DUMAS MILNE EDWARDS,SEVELIN JOBERT,JEAN
                GIORDANO
            PC C12N15/09,C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12.
                C12N1/21,
            PC C12N5/10,C12P21/02,C12P21/08,C12Q1/68,C12N15/00,C12N
                C12N15/00
            CC EST and encoded human protein
            FH Key   Location/Qualifiers
            FT CDS   341..517.
FEATURES   Location/Qualifiers
     source 1..522
             /organism="Homo sapiens"
             /mol_type="genomic DNA"
             /db_xref="taxon:9606"
ORIGIN
Query Match   1.9%; Score 26; DB 6; Length 522;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 26; Conservative 0; Mismatches 0; Indels 0;

Qy   1 ATGTCATTGTTAGACTTTGAAATTTC 26
          ||||||||||||||||||||||||||
Db 476 ATGTCATTGTTAGACTTTGAAATTTC 501

RESULT 59
AX381939
LOCUS   AX381939   531 bp   DNA   linear   PAT 1
DEFINITION Sequence 877 from Patent WO0212280.
ACCESSION   AX381939
VERSION   AX381939.1 GI:19576761
KEYWORDS
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutel
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
            1
REFERENCE   Pyle,R.A., Xu,J. and Secrist,H.
AUTHORS   Compositions and methods for the therapy and diagnosis of
TITLE   cancer
            Patent: WO 0212280-A 877 14-FEB-2002;
JOURNAL   CORIXA CORPORATION (US)
FEATURES   Location/Qualifiers
     source 1..531
             /organism="Homo sapiens"
             /mol_type="unassigned DNA"
             /db_xref="taxon:9606"
ORIGIN
Query Match   1.9%; Score 26; DB 6; Length 531;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 26; Conservative 0; Mismatches 0; Indels 0;

Qy   1 ATGTCATTGTTAGACTTTGAAATTTC 26
          ||||||||||||||||||||||||||
Db 459 ATGTCATTGTTAGACTTTGAAATTTC 484

---

apiens (human)
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ases 1 to 516)
s,J.B.D.M., Duclair,E. and Jordan,J.Y.
ce tag and encoded human protein
:: JP 2001269182-A 1263 02-OCT-2001;
Homo sapiens (human)
JP 2001269182-A/1263
02-OCT-2001
24-FEB-2000 JP 2000118773
26-FEB-1999 US   60/122487
JEAN BAPUTIST DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
JORDAN
C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
5/10,
C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
F15/40
Key   Location/Qualifiers
CDS   343..516.
                Location/Qualifiers
        1..516
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"

larity 100.0%; Pred. No. 0.022; Indels 0; Gaps 0;
Conservative 0; Mismatches 0;

TCATTGTTAGACTTTGAAATTTC 26
||||||||||||||||||||||||||
TCATTGTTAGACTTTGAAATTTC 503

                 522 bp   DNA   linear   PAT 18-DEC-2003
934
nce 3571 from patent US 6639063.
934.1 GI:40171044
wn.
wn.
ssified.
ds,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
and encoded human proteins
t: US 6639063-A 3571 28-OCT-2003;
Location/Qualifiers
        1..522
        /organism="unknown"
        /mol_type="genomic DNA"

larity 100.0%; Pred. No. 0.022; Indels 0; Gaps 0;
Conservative 0; Mismatches 0;

TCATTGTTAGACTTTGAAATTTC 26
||||||||||||||||||||||||||
TCATTGTTAGACTTTGAAATTTC 501

                 522 bp   DNA   linear   PAT 18-SEP-2002
487
nd encoded human protein.
487

22   534 bp   mRNA   linear   PRI 14-NOV-2003
apiens cDNA clone IMAGE:3839189, partial cds.
22
22.1   GI:38328364

apiens (human)
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 534)
berg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
er,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
ul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
s,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
enko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
ton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
z,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
ci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
on,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
an,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
on,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
z,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
rd,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
n,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
field,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
ch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
tion and initial analysis of more than 15,000 full-length
and mouse cDNA sequences
Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
57
32
ses 1 to 534)
berg,R.
Submission
ted (13-NOV-2003) National Institutes of Health, Mammalian
ollection (MGC), Cancer Genomics Office, National Cancer
ute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

C Project URL: http://mgc.nci.nih.gov
t: MGC help desk
    cgapbs-r@mail.nih.gov
Procurement: ATCC
ibrary Preparation: CLONTECH Laboratories, Inc.
ibrary Arrayed by: The I.M.A.G.E. Consortium (LLNL)
quencing by: Genome Sequence Centre,
cer Agency, Vancouver, BC, Canada
cgsc.bc.ca
Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
a Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
ia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
e, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
, Thorne, Miranda Tsai, Natasja van den Bosch, Jill Vardy,
Yang, Scott Zuyderduyn, Marco Marra.

distribution: MGC clone distribution information can be found
h the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
:: IRAL Plate: 51 Row: c Column: 21
clone was selected for full length sequencing because it
t the following selection criteria: Hexamer frequency ORF
i8.

Location/Qualifiers
1..534
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="IMAGE:3839189"

                /tissue_type="Brain, glioblastoma"
                /clone_lib="NIH_MGC_57"
                /lab_host="DH10B"
                /note="Vector: pDNR-LIB"

ORIGIN

Query Match   1.9%;   Score 26;   DB 9; Length 534;
Best Local Similarity   100.0%;   Pred. No. 0.022;
Matches 26; Conservative   0; Mismatches   0; Indels   0;   C

Qy   1 ATGTCATTGTTAGACTTTGAAATTTC 26
        |||||||||||||||||||||||||||
Db 438 ATGTCATTGTTAGACTTTGAAATTTC 463

RESULT 61
AR415931
LOCUS   AR415931   538 bp   DNA   linear   PAT 1E
DEFINITION Sequence 3568 from patent US 6639063.
ACCESSION   AR415931
VERSION   AR415931.1 GI:40171041
KEYWORDS   Unknown.
SOURCE   Unknown.
ORGANISM   Unclassified.
REFERENCE   1 (bases 1 to 538)
AUTHORS   Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
TITLE   EST's and encoded human proteins
JOURNAL   Patent: US 6639063-A 3568 28-OCT-2003;
FEATURES   Location/Qualifiers
     source   1..538
                /organism="unknown"
                /mol_type="genomic DNA"

ORIGIN

Query Match   1.9%;   Score 26;   DB 6; Length 538;
Best Local Similarity   100.0%;   Pred. No. 0.022;
Matches 26; Conservative   0; Mismatches   0; Indels   0;   C

Qy   1 ATGTCATTGTTAGACTTTGAAATTTC 26
        |||||||||||||||||||||||||||
Db 451 ATGTCATTGTTAGACTTTGAAATTTC 476

RESULT 62
BD111484
LOCUS   BD111484   538 bp   DNA   linear   PAT 1E
DEFINITION EST and encoded human protein.
ACCESSION   BD111484
VERSION   BD111484.1 GI:23206302
KEYWORDS   JP 2002010789-A/3561.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 538)
AUTHORS   Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
TITLE   EST and encoded human protein
JOURNAL   Patent: JP 2002010789-A 3561 15-JAN-2002;
            GENSET CORP
COMMENT   OS   Homo sapiens (human)
            PN   JP 2002010789-A/3561
            PD   15-JAN-2002
            PF   07-AUG-2000 JP 2000280989
            PR   05-AUG-1999 US   60/147499
            PI   JEAN BAPTIST DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN
                  GIORDANO
            PC   C12N15/09,C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N
                  C12N1/21,
            PC   C12N5/10,C12P21/02,C12P21/08,C12Q1/68,C12N15/00,C12NE
                  C12N15/00
            CC   EST and encoded human protein
            FH   Key   Location/Qualifiers

CDS   316..492.
                Location/Qualifiers
                1..538
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"

larity   1.9%; Score 26; DB 6; Length 538;
Conservative   0; Mismatches 0; Indels 0; Gaps 0;

TCATTGTTAGACTTTGAAATTTC 26
||||||||||||||||||||||
TCATTGTTAGACTTTGAAATTTC 476

406   540 bp   DNA   linear   PAT 18-DEC-2003
nce 1269 from Patent EP1033401.
406
406.1 GI:40041535
sapiens (human)
sapiens
yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
lia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
ssed sequence tags and encoded human proteins
t: EP 1033401-A 1269 06-SEP-2000;
t (FR)
                Location/Qualifiers
                1..540
                /organism="Homo sapiens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:9606"
                367..>540
                /note="unnamed protein product"
                /codon_start=1
                /protein_id="CAB99544.1"
                /db_xref="GI:40041536"
                /translation="MEYERSPLERKGQTLCFHSEDIAEPVPQGYCIHSLSLKGCATS
                RMSLLDFEISNCLP"

larity   1.9%; Score 26; Pred. No. 0.022; Length 540;
Conservative   0; Mismatches 0; Indels 0; Gaps 0;

TCATTGTTAGACTTTGAAATTTC 26
|||||||||||||||||||||
TCATTGTTAGACTTTGAAATTTC 527

016   540 bp   DNA   linear   PAT 27-AUG-2002
nce tag and encoded human protein.
016
016.1 GI:22556239
01269182-A/1262.
sapiens (human)
sapiens
yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
lia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ds,J.B.D.M., Duclair,E. and Jordan,J.Y.
nce tag and encoded human protein
t: JP 2001269182-A 1262 02-OCT-2001;
T
Homo sapiens (human)

PN   JP 2001269182-A/1262
PD   02-OCT-2001
PF   24-FEB-2000 JP 2000118773
PR   26-FEB-1999 US 60/122487
PI   JEAN BAPUTIST DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JE
PI   JORDAN
PC   C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12
     C12N5/10,
PC   C12P21/02,C12P21/08,C12P21/08//G06F17/30,C12N15/00,C1
     G06F15/40
CC
FH   Key   Location/Qualifiers
FT   CDS   367..540.
FEATURES   Location/Qualifiers
     source   1..540
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match   1.9%; Score 26; DB 6; Length 540;
Best Local Similarity 100.0%; Pred. No. 0.022;   Indels 0;
Matches   26; Conservative   0; Mismatches 0;

Qy   1 ATGTCATTGTTAGACTTTGAAATTTC 26
           ||||||||||||||||||||||||||
Db   502 ATGTCATTGTTAGACTTTGAAATTTC 527

RESULT 65
BD224094
LOCUS   BD224094   540 bp   DNA   linear   PAT 1
DEFINITION   Sequence characteristic to gene transcription controlled
             hypoxia.
ACCESSION   BD224094
VERSION   BD224094.1 GI:33033864
KEYWORDS   JP 2002525081-A/9.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutel
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
REFERENCE   1 (bases 1 to 540)
AUTHORS   Einat,P., Skaliter,R. and Feinstein,E.
TITLE   Sequence characteristic to gene transcription controlled
JOURNAL   Patent: JP 2002525081-A 9 13-AUG-2002;
             QUARK BIOTECH INC
COMMENT   OS   Homo sapiens (human)
             PN   JP 2002525081-A/9
             PD   13-AUG-2002
             PF   27-AUG-1999 JP 2000571058
             PR   27-AUG-1998 US 60/098158, 05-MAY-1999 US 60/132
             PAZ EINAT,RAMI SKALITER,ELENA FEINSTEIN
             PC   C12N15/09,A61K31/711,A61K45/00,A61P9/00,A6
                  A61P43/00,
             PC   C12Q1/68,G01N33/53,G01N33/566,G01N33/58,G01N33/68,G0
             PC   C12N15/00
             CC   Sequence characteristic to gene transcription contro
             CC   hypoxia
             FH   Key   Location/Qualifiers
             FT   source   1..540
                                  /organism='Homo sapiens (human)' .
FEATURES   Location/Qualifiers
     source   1..540
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match   1.9%; Score 26; DB 6; Length 540;
Best Local Similarity 100.0%; Pred. No. 0.022;   Indels 0;
Matches   26; Conservative   0; Mismatches 0;

Qy   1 ATGTCATTGTTAGACTTTGAAATTTC 26

```
Matches  26;  Conservative   0;  Mismatches   0;  Indels   0;  C

Qy     1 ATGTCATTGTTAGACTTTGAAATTTC 26
         ||||||||||||||||||||||||||
Db   473 ATGTCATTGTTAGACTTTGAAATTTC 498
```

RESULT 68
BC042807

```
LOCUS       BC042807                 560 bp    mRNA    linear   PRI 0?
DEFINITION  Homo sapiens, clone IMAGE:5285034, mRNA.
ACCESSION   BC042807
VERSION     BC042807.1  GI:27502998
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 560)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JAN-2003) National Institutes of Health, Man
            Gene Collection (MGC), Cancer Genomics Office, National Ca
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 2089?
            USA
REMARK      NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI) &
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louise
            Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanav
            A.N., Gibbs, R.A.

            Clone distribution: MGC clone distribution information car
            through the I.M.A.G.E. Consortium/LLNL at: http://image.ll
            Series: IRAK Plate: 91 Row: g Column: 13.
FEATURES             Location/Qualifiers
     source          1..560
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5285034"
                     /tissue_type="Brain, hippocampus"
                     /clone_lib="NIH_MGC_95"
                     /lab_host="DH10B"
                     /note="Vector: pBluescript"
ORIGIN

Query Match        1.9%;  Score 26;  DB 9;  Length 560;
Best Local Similarity  100.0%;  Pred. No. 0.022;
Matches  26;  Conservative   0;  Mismatches   0;  Indels   0;  C

Qy     1 ATGTCATTGTTAGACTTTGAAATTTC 26
         ||||||||||||||||||||||||||
Db   478 ATGTCATTGTTAGACTTTGAAATTTC 503
```

RESULT 69
AR415932

```
LOCUS       AR415932                 595 bp    DNA     linear   PAT 1?
DEFINITION  Sequence 3569 from patent US 6639063.
ACCESSION   AR415932
VERSION     AR415932.1  GI:40171042
KEYWORDS    .
SOURCE      Unknown.
```

```
             |||||||||||||||||||||||||
          CATTGTTAGACTTTGAAATTTC 512

                            555 bp    DNA     linear   PAT 18-DEC-2003
      ce 3572 from patent US 6639063.
35
35.1  GI:40171045
n.
n.
sified.
ses 1 to 555)
s,J.-B.D.M.,Jobert,S. and Giordano,J.-Y.
 and encoded human proteins
: US 6639063-A 3572 28-OCT-2003;
             Location/Qualifiers
             1..555
                     /organism="unknown"
                     /mol_type="genomic DNA"

arity      1.9%;  Score 26;  DB 6;  Length 555;
onservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

      CATTGTTAGACTTTGAAATTTC 26
      ||||||||||||||||||||||
      CATTGTTAGACTTTGAAATTTC 498

88                          555 bp    DNA     linear   PAT 18-SEP-2002
d encoded human protein.
88
88.1  GI:23206306
2010789-A/3565.
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 555)
s,J.B.D.M., Jobert,S. and Giordano,J.E.
d encoded human protein
: JP 2002010789-A 3565 15-JAN-2002;
   CORP
omo sapiens (human)
P 2002010789-A/3565
5-JAN-2002
7-AUG-2000 JP 2000280989
5-AUG-1999 US   60/147499
EAN BAPTIST DUMAS MILNE EDWARDS,SEVELIN JOBERT,JEAN EVE  PI
ANO
12N15/09,C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19, PC
/21,
12N5/10,C12P21/02,C12P21/08,C12Q1/68,C12N15/00,C12N5/00, PC
/00
=a, g, c or t      Location/Qualifiers
ey          338..514
DS          338..514
isc_feature 541.
             Location/Qualifiers
             1..555
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

arity      1.9%;  Score 26;  DB 6;  Length 555;  Pred. No. 0.022;
```

```
VERSION     AK130060.1  GI:34526798
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutel
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
REFERENCE   1
  AUTHORS   Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura
            Fukuzumi,Y., Fujimori,Y., Komiyama,M., Suzuki,Y., Hata,H.
            Nakagawa,K., Mizuno,S., Morinaga,M., Kawamura,M., Sugiyam
            Irie,R., Otsuki,T., Sato,H., Nishikawa,T., Sugiyama,A.,
            Kawakami,B., Nagai,K., Isogai,T. and Sugano,S.
            NEDO human cDNA sequencing project
  TITLE     Unpublished
  JOURNAL
REFERENCE   2  (bases 1 to 1822)
  AUTHORS   Sugano,S. and Suzuki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JUL-2003) Sumio Sugano, Institute of Medica
            University of Tokyo, Laboratory of Genome Structure, Huma
            Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
            (E-mail:flcdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
            Fax:81-3-5449-5416)
COMMENT     NEDO human cDNA sequencing project supported by Ministry
            Economy, Trade and Industry of Japan; cDNA full insert se
            Research Association for Biotechnology (RAB); cDNA librar
            construction and 5'-end one pass sequencing: Institute of
            Science, University of Tokyo, Laboratory of Genome Struct
            Genome Center; 3'-end one pass sequencing: RAB; clone sel
            full insert sequencing: RAB and Helix Research Institute.
FEATURES    Location/Qualifiers
     source          1..1822
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="LNF01586"
                     /cell_type="fibroblast"
                     /tissue_type="lung"
                     /clone_lib="LNF"
                     /note="cloning vector: pME18SFL3"
ORIGIN
Query Match             1.9%;  Score 26;  DB 9;  Length 1822;
Best Local Similarity   100.0%;  Pred. No. 0.025;  Indels  0;
Matches   26;  Conservative   0;  Mismatches   0;

Qy          1 ATGTCATTGTTAGACTTTGAAATTTC 26
              ||||||||||||||||||||||||||
Db       1756 ATGTCATTGTTAGACTTTGAAATTTC 1781
```

```
RESULT 72
HSM802687
LOCUS       HSM802687               2183 bp    mRNA    linear   PRI 1
DEFINITION  Homo sapiens mRNA; cDNA DKFZp762B195 (from clone DKFZp762
ACCESSION   AL359585
VERSION     AL359585.1  GI:8655645
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutel
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
REFERENCE   1  (bases 1 to 2183)
  AUTHORS   Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Weil,B. a
            Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUN-2000) MIPS, Am Klopferspitz 18a, D-8215
            Martinsried, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German
            Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.c
            sequenced by GBF (National Research Centre for Biotechnol
            Braunschweig/Germany) within the cDNA sequencing consorti
            German Genome Project.
            This clone (DKFZp762B195) is available at the RZPD in Ber
```

```
wn.
ssified.
ases 1 to 595)
is,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
and encoded human proteins
:: US 6639063-A 3569 28-OCT-2003;
     Location/Qualifiers
     1..595
                     /organism="unknown"
                     /mol_type="genomic DNA"
larity       1.9%;  Score 26;  DB 6;  Length 595;
100.0%;  Pred. No. 0.022;  Indels  0;  Gaps  0;
Conservative   0;  Mismatches   0;

TCATTGTTAGACTTTGAAATTTC 26
|||||||||||||||||||||||
TCATTGTTAGACTTTGAAATTTC 538
```

```
485                     595 bp    DNA     linear       PAT 18-SEP-2002
nd encoded human protein.
485.1  GI:23206303
02010789-A/3562.
sapiens (human)
yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
lia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ases 1 to 595)
ds,J.B.D.M., Jobert,S. and Giordano,J.E.
nd encoded human protein
t: JP 202010789-A 3562 15-JAN-2002;
T CORP
Homo sapiens (human)
JP 202010789-A/3562
15-JAN-2002
07-AUG-2000 JP 2000280989
05-AUG-1999 US 60/147499
JEAN BAPUTIST DUMAS MILNE EDWARDS,SEVELIN JOBERT,JEAN EVE  PI
DANO
C12N15/09,C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19, PC
1/21,
C12N5/10,C12P21/02,C12P21/08,C12Q1/68,C12N15/00,C12N5/00, PC
5/00
EST and encoded human protein
Key             Location/Qualifiers
CDS             378..554.
     Location/Qualifiers
     1..595
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
larity       1.9%;  Score 26;  DB 6;  Length 595;
100.0%;  Pred. No. 0.022;  Indels  0;  Gaps  0;
Conservative   0;  Mismatches   0;

TCATTGTTAGACTTTGAAATTTC 26
|||||||||||||||||||||||
TCATTGTTAGACTTTGAAATTTC 538
```

```
060                     1822 bp   mRNA    linear       PRI 10-SEP-2003
sapiens cDNA FLJ26550 fis, clone LNF01586, highly similar to
aldolase (EC 2.2.1.2).
060
```

```
ORIGIN
  Query Match
  Best Local Similarity   1.9%;   Score 26;   DB 9;   Length 2650;
  Matches 26;  Conservative   0;  Mismatches  0;   Indels  0;   G

  Qy     1 ATGTCATTGTTAGACTTTGAAATTC 26
           |||||||||||||||||||||||||
  Db  2584 ATGTCATTGTTAGACTTTGAAATTC 2609

RESULT 74
HSDJ686N3
LOCUS        HSDJ686N3        110293 bp    DNA    linear    PRI 06
DEFINITION   Human DNA sequence from clone RP4-686N3 on chromosome 20q1
             Contains the 3' part of the gene for a novel ATP dependent
             helicase (contains conserved C-terminal helicase domains a
             DEAD/DEAH boxes), the KIAA1404 gene, a putative novel gene
             STSs, GSSs and two CpG islands, complete sequence.
ACCESSION    AL049766
VERSION      AL049766.14  GI:5763746
KEYWORDS     HTG; CpG island; DEAD box; DEAH box; KIAA1404; RNA helicas
SOURCE       Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 110293)
  AUTHORS    Corby,N.
  TITLE      Direct Submission
  JOURNAL    Submitted (01-MAR-2001) Sanger Centre, Hinxton, Cambridges
             CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clon
             requests: clonerequest@sanger.ac.uk
COMMENT      On Aug 24, 1999 this sequence version replaced gi:5730221.
             During sequence assembly data is compared from overlapping
             Where differences are found these are annotated as variati
             together with a note of the overlapping clone name. Note t
             variation annotation may not be found in the sequence subm
             corresponding to the overlapping clone, as we submit seque
             only a small overlap as described above.
             The following abbreviations are used to associate primary
             numbers given in the feature table with their source datab
             Em., EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Info
             on the WORMPEP database can be found at
             http://www.sanger.ac.uk/Projects/C_elegans/wormpep This se
             the entire insert of clone RP4-686N3 This sequence was fin
             follows unless otherwise noted: all regions were either
             double-stranded or sequenced with an alternate chemistry o
             by high quality data (i.e. phred quality >= 30); an attem
             made to resolve all sequencing problems, such as compressi
             repeats; all regions were covered by at least one plasmid
             or more than one M13 subclone; and the assembly was confir
             restriction digest. This sequence was generated from part
             bacterial clone contigs of human chromosome 20, constructe
             Sanger Centre Chromosome 20 Mapping Group. Further inform
             be found at http://www.sanger.ac.uk/HGP/Chr20
             RP4-686N3 is from the library RPCI-4 constructed by the gr
             Pieter de Jong. For further details see
             http://www.chori.org/bacpac/home.htm
             VECTOR: pCYPAC2.
FEATURES     Location/Qualifiers
     source  1..110293
             /organism="Homo sapiens"
             /mol_type="genomic DNA"
             /db_xref="taxon:9606"
             /chromosome="20"
             /map="q13.2-13.2"
             /clone="RP4-686N3"
             /clone_lib="RPCI-4"
     repeat_region  1..61
             /note="AluY repeat: matches 2..62 of consensus"
             207..21327
     gene    /gene="dJ686N3.1"
     mRNA    join(<207..266,518..683,2195..2241,2363..2449
```

```
contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
ation about the clone and the sequencing project is available
://www.mips.biochem.mpg.de/proj/cDNA/.
             Location/Qualifiers
             1..2183
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /map="20q13.2-13.2"
             /clone="DKF2p762B195"
             /tissue_type="melanoma (MeWo cell line)"
             /clone_lib="762 (synonym: hmel2). Vector pSport1; host
             DH10B; sites NotI + SalI"
             /dev_stage="adult"
             2140..2145
             2152

arity    1.9%;   Score 26;   DB 9;   Length 2183;
onservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

CATTGTTAGACTTTGAAATTC 26
|||||||||||||||||||||
CATTGTTAGACTTTGAAATTC 2116

75                    2650 bp    mRNA    linear    PRI 09-SEP-2003
piens cDNA FLJ42181 fis, clone THYMU2031368.
75
75.1  GI:34529902
capping; fis (full insert sequence).
piens (human)
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
la; Eutheria; Primates; Catarrhini; Hominidae; Homo.

.A., Takahashi-Fujii,A., Tanase,T., Imose,N., Takeuchi,K.,
M., Musashino,K., Yuuki,H., Hara,H., Sugiyama,T., Irie,R.,
T., Sato,H., Ota,T., Wakamatsu,A., Ishii,S., Yamamoto,J.,
., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K.,
ta,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H.,
C., Wagatsuma,M., Murakawa,K., Kanehori,K., Sugiyama,A.,
ni,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K.
ogai,T.
aman cDNA sequencing project
ished
es 1 to 2650)
T. and Yamamoto,J.
Submission
ed (15-JUL-2003) Takao Isogai, FLJ Project(HRI Team); 2-6-7
-Kamatari, Kisarazu, Chiba 292-0818, Japan
L:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
man cDNA sequencing project supported by Ministry of
/, Trade and Industry of Japan; cDNA full insert sequencing:
ch Association for Biotechnology (RAB); cDNA library
ction: Helix Research Institute (HRI) (supported by Japan
chnology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
d Biotechnology Center, National Institute of Technology and
cion; clone selection for full insert sequencing: HRI and
nnotation: HRI and RAB.
             Location/Qualifiers
             1..2650
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="THYMU2031368"
             /tissue_type="thymus"
             /clone_lib="THYMU2"
             /note="cloning vector: pME18SFL3"
```

```
5965. .6138,7449. .7599,10556. .10655,10818. .10959,
12185. .12338,13400. .13481,13583. .13760,16199. .16305,
16486. .16588,19143. .19237,19333. .19456,19844. .19931,
20879. .20953,21066. .21327)
/gene="dJ686N3.1"
/product="dJ686N3.1 (novel ATP dependent RNA helicase
(contains conserved C-terminal helicase domains and
DEAD/DEAH boxes))"
/evidence=not experimental
join(<207. .266,518. .683,2195. .2241,2363. .2449,3680. .3785,
5965. .6138,7449. .7599,10556. .10655,10818. .10959,
12185. .12338,13400. .13481,13583. .13760,16199. .16305,
16486. .16588,19143. .19237,19333. .19456,19844. .19931,
20879. .20953,21066. .21084)
/gene="dJ686N3.1"
/note="continues from bA120I11 (AL357560)
match: proteins: Tr:O23506 Tr:Q9ZI52 Tr:Q92CQ0 Wp:CE00337
Tr:Q9ZR28 Sw:P25888 Sw:P32892 Sw:Q09903 Tr:O26305
Tr:O87951 Wp:CE22901 Tr:Q07886 Tr:O07897 Tr:O25029
Sw:P38712 Tr:Q9ZMI8 Tr:Q92Z6C9 Sw:P21693 Sw:P44701"
/codon_start=1
/evidence=not experimental
/product="dJ686N3.1 (novel ATP dependent RNA helicase
(contains conserved C-terminal helicase domains and
DEAD/DEAH boxes))"
/protein_id="CAB65627.2"
/db_xref="GI:13274192"
/db_xref="GOA:Q96GQ7"
/db_xref="SWISS-PROT:Q96GQ7"
/translation="RAATLDEKIEKVRKKRKTEDKEAKSGKLEKEKEAKEGSEPKEQ
EDLQENDEBGSEDEASETDYSSADENILTKADTLKVKDRKKKKKKGQEAGGFFEDASQ
YDBNLSFQDMNLSRPLLKAITAMGFKQPTPIQKACIPVGLLGKDICACAATGTKTAA
FALPVLERLIYKPRQAPVTRVLVLVPTRELGIQVHSVTRQLAQFCNITTCLAVGGLDV
KSQEAALRAAPDILIATPGRLIDHLHNCPSFHLSSIEVLILDEADRMLDEYFEEQMKE
IIRMCSHHRQTMLFSAIMTDEVKDLASVSLKNPVRIFVNSNTDVAPFLRQEFIRIRPN
REGDREAIVAALLTRTFDHVMLFTQTKQAHRMHILLGLQVGELHGNLSQTQRL
EALRRFKDEQIDILVATDVAARGLDIEGVKTVINFTMPNTIKHYVHRVGRTARAGRAG
RSVSLVGEDERKMLKEIVKAAKAPVKARILPQDVILKFRDKIEKMEKDVYAVLQLEAE
EKEMQQSEAQINTAKRLLEKGKEAVVQEPERSWFQTKBERKEKIAKALQEFDLALRG
KKKRKRKFMKDAKKKGEMTAEERSQFEILKAQMPAERLAKRNRRAKRARAMPEEEPVRG
PAKKQKQGKKSVFDEELTNTSKKALKQYRAGPSFEERRKQLGLPHQRRGGNFKSKSRYK
RRK"

c  849. .1146
   /note="AluSx repeat: matches 1. .298 of consensus"
c  1147. .1423
   /note="AluJo repeat: matches 4. .282 of consensus"
c  1605. .1912
   /note="AluSx repeat: matches 1. .307 of consensus"
   complement(1913. .2334)
c  1952. .2099
   /note="match: GSS: Em:AQ343240"
c  2473. .2589
   /note="MIR repeat: matches 1. .170 of consensus"
c  2617. .2703
   /note="FLAM_C repeat: matches 2. .125 of consensus"
c  2704. .3004
   /note="L2 repeat: matches 2051. .2135 of consensus"
c  3005. .3572
   /note="AluY repeat: matches 1. .301 of consensus"
c  3908. .4051
   /note="L2 repeat: matches 2135. .2708 of consensus"
c  4108. .4411
   /note="FLAM_C repeat: matches 1. .133 of consensus"
n  4441. .4610
   /note="AluYa5 repeat: matches 1. .304 of consensus
   AluYa5 repeat: matches 1. .304 of consensus"
n  4624. .4917
   /note="L1MD3 repeat: matches 7557. .7740 of consensus"
n  5089. .5388
   /note="AluSx repeat: matches 2. .293 of consensus"
n  5392. .5685
   /note="AluSg repeat: matches 1. .300 of consensus"
   /note="AluSq repeat: matches 1. .301 of consensus"
```

```
repeat_region  5690. .5809
   /note="AluSq/x repeat: matches 14. .133 of conse
   AluSq/x repeat: matches 14. .133 of consensus"
repeat_region  6281. .6579
   /note="AluY repeat: matches 1. .303 of consensus
repeat_region  6896. .7199
   /note="AluSg1 repeat: matches 1. .300 of consens
   AluSx repeat: matches 1. .303 of consensus"
repeat_region  7814. .7931
   /note="MIR repeat: matches 48. .174 of consensus
repeat_region  8181. .8293
   /note="L2 repeat: matches 1829. .1957 of consens
repeat_region  8294. .8596
   /note="L2 repeat: matches 2347. .2704 of consens
repeat_region  8870. .9009
   /note="L1MC4 repeat: matches 7203. .7351 of cons
repeat_region  9126. .9436
   /note="AluJo repeat: matches 1. .308 of consensu
repeat_region  9437. .9492
   /note="AluJo repeat: matches 86. .141 of consens
repeat_region  9495. .9597
   /note="L1MC4 repeat: matches 7525. .7624 of cons
repeat_region  9598. .9731
   /note="AluSp repeat: matches 1. .127 of consensu
repeat_region  9732. .10040
   /note="AluY repeat: matches 1. .308 of consensus
repeat_region  10041. .10191
   /note="AluSp repeat: matches 127. .300 of consen
repeat_region  10192. .10371
   /note="L1MC4 repeat: matches 7624. .7814 of cons
repeat_region  11182. .11513
   /note="L2 repeat: matches 1988. .2346 of consens
repeat_region  11514. .11813
   /note="AluSx repeat: matches 1. .300 of consensu
repeat_region  11814. .11963
   /note="L2 repeat: matches 2346. .2489 of consens
repeat_region  12695. .12745
   /note="L2 repeat: matches 2648. .2700 of consens
repeat_region  12703. .12773
   /note="MIR repeat: matches 178. .250 of consensu
repeat_region  12997. .13292
   /note="AluY repeat: matches 1. .296 of consensus
misc_feature  complement(13667. .13816)
   /note="match: GSS: Em:AQ413681"
misc_feature  complement(13759. .14174)
   /note="match: GSS: Em:AQ252228"
repeat_region  13817. .13960
   /note="match: GSS: Em:AQ252228"
misc_feature  13961. .14303
   /note="AluSx repeat: matches 8. .155 of consensus"
   /gene="dJ686N3.1"
misc_feature  13961. .14174
   /gene="dJ686N3.1"
repeat_region  14175. .14262
   /note="AluJ/FLAM repeat: matches 2. .87 of conse
   Alu/FLAM repeat: matches 2. .87 of consensus"
repeat_region  14304. .14596
   /note="AluJo repeat: matches 3. .300 of consensu
repeat_region  14744. .15145
   /note="L2 repeat: matches 2251. .2749 of consens
repeat_region  15261. .15558
   /note="AluJo repeat: matches 9. .306 of consensu
repeat_region  15560. .15717
   /note="AluJo repeat: matches 145. .293 of consen
repeat_region  15718. .16016
   /note="AluSx repeat: matches 1. .297 of consensus
```

```
Query Match     1.9%;    Score 26;    DB 9;    Length 110293;
Best Local Similarity   100.0%;    Pred. No. 0.035;
Matches   26;    Conservative   0;    Mismatches   0;    Indels   0;

Qy     1  ATGTCATTGTTAGACTTTGAAATTTC  26
```

```
|||||||||||||||||||||||||
CATTGTAGACTTGAAATTTC 66465
```

```
                 159020 bp    DNA        linear   HTG 27-APR-2000
23
piens chromosome 20 clone RP11-564F22 map 20, WORKING DRAFT
E, 21 unordered pieces.
23.2  GI:7651931
GS PHASE1; HTGS_DRAFT.
piens (human)
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 159020)
B., Linton,L., Nusbaum,C. and Lander,E.
piens chromosome 20, clone RP11-564F22
shed
ses 1 to 159020)
B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
on,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
vkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
ano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
ore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
,Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
,Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
ierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
,LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
y,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
 Santos,R., Schauer,S., Severy,P., Spencer,B.,
Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
ev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
,Zainoun,J., Zimmer,A. and Zody,M.
 Submission
ed (25-MAR-2000) Whitehead Institute/MIT Center for Genome
h, 320 Charles Street, Cambridge, MA 02141, USA
27, 2000 this sequence version replaced gi:7328901.
eats were identified using RepeatMasker:
'F.A. & Green,P. (1996-1997)
'ftp.genome.washington.edu/RM/RepeatMasker.html
---- Genome Center
uter: Whitehead Institute/ MIT Center for Genome Research
ter code: WIBR
 site: http://www-seq.wi.mit.edu
tact: sequence submissions@genome.wi.mit.edu
----- Project Information
ter project name: L8824
ter clone name: 564 F 22
---- Summary Statistics
quencing vector: M13; M77815; 100% of reads
emistry: Dye-terminator Big Dye; 100% of reads
embly program: Phrap; version 0.960731
isensus quality: 147720 bases at least Q40
isensus quality: 153269 bases at least Q30
isensus quality: 155491 bases at least Q20
sert size: 171000; agarose-fp
sert size: 157020; sum-of-contigs
lity coverage: 3.9 in Q20 bases; agarose-fp
lity coverage: 4.2 in Q20 bases; sum-of-contigs
: This is a 'working draft' sequence. It currently
sts of 21 contigs. The true order of the pieces
ot known and their order in this sequence record is
rary. Gaps between the contigs are represented as
```

```
*  runs of N, but the exact sizes of the gaps are unknown.
*  This record will be updated with the finished sequence
*  as soon as it is available and the accession number will
*  be preserved.
*
*        1    1127: contig of 1127 bp in length
*     1128    1227: gap of 100 bp
*     1228    2315: contig of 1088 bp in length
*     2316    2415: gap of 100 bp
*     2416    4526: contig of 2111 bp in length
*     4527    4626: gap of 100 bp
*     4627    8382: contig of 3756 bp in length
*     8383    8482: gap of 100 bp
*     8483   12056: contig of 3574 bp in length
*    12057   12156: gap of 100 bp
*    12157   14923: contig of 2767 bp in length
*    14924   15023: gap of 100 bp
*    15024   19515: contig of 4492 bp in length
*    19516   19615: gap of 100 bp
*    19616   23997: contig of 4382 bp in length
*    23998   24097: gap of 100 bp
*    24098   27747: contig of 3650 bp in length
*    27748   27847: gap of 100 bp
*    27848   33472: contig of 5625 bp in length
*    33473   33572: gap of 100 bp
*    33573   40393: contig of 6821 bp in length
*    40394   40493: gap of 100 bp
*    40494   47274: contig of 6781 bp in length
*    47275   47374: gap of 100 bp
*    47375   53409: contig of 6035 bp in length
*    53410   53509: gap of 100 bp
*    53510   62810: contig of 9301 bp in length
*    62811   62910: gap of 100 bp
*    62911   71685: contig of 8775 bp in length
*    71686   71785: gap of 100 bp
*    71786   80849: contig of 9064 bp in length
*    80850   80949: gap of 100 bp
*    80950   90631: contig of 9682 bp in length
*    90632   90731: gap of 100 bp
*    90732   98024: contig of 7293 bp in length
*    98025   98124: gap of 100 bp
*    98125  110319: contig of 12195 bp in length
*   110320  110419: gap of 100 bp
*   110420  127336: contig of 16917 bp in length
*   127337  127436: gap of 100 bp
*   127437  159020: contig of 31584 bp in length.

FEATURES             Location/Qualifiers
     source          1..159020
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="20"
                     /clone="RP11-564F22"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1..1127
                     /note="assembly_fragment"
     misc_feature    1228..2315
                     /note="assembly_fragment"
     misc_feature    2416..4526
                     /note="assembly_fragment"
     misc_feature    4627..8382
                     /note="assembly_fragment"
     misc_feature    8483..12056
                     /note="assembly_fragment"
     misc_feature    12157..14923
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
     misc_feature    15024..19515
                     /note="assembly_fragment"
     misc_feature    19616..23997
                     /note="assembly_fragment"
     misc_feature    24098..27747
```

```
27848..33472
/note="assembly_fragment"
33573..40393
/note="assembly_fragment"
40494..47274
/note="assembly_fragment"
47375..53409
/note="assembly_fragment"
53510..62810
/note="assembly_fragment"
62911..71685
/note="assembly_fragment"
71786..80849
/note="assembly_fragment"
80950..90631
/note="assembly_fragment"
90732..98024
/note="assembly_fragment
clone_end:SP6
vector_side:right"
98125..110319
/note="assembly_fragment"
110420..127336
/note="assembly_fragment"
127437..159020
/note="assembly_fragment"
```

```
Similarity   1.9%;  Score 26;  DB 2;  Length 159020;
Pred. No. 0.036;
Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

CCATTGTTAGACTTTGAAATTTC 26
||||||||||||||||||||||
CCATTGTTAGACTTTGAAATTTC 93106
```

```
April 8, 2004, 22:35:57
pcs
```

GenCore version 5.1.6
opyright (c) 1993 - 2004 Compugen Ltd.

c search, using sw model

il 8, 2004, 20:38:11 ; Search time 3960 Seconds
                         (without alignments)
             10353.731 Million cell updates/sec

09-245-198A-3
3
tgtcattgtagacttga.........gacaaaatgttgataaatgg 1373

GO_NUC
op 60.0 , Gapext 60.0

13289 seqs, 14931090276 residues

s satisfying chosen parameters: 55025678

th: 0
th: 2000000000

sting first 100 summaries

T:*
em_estba:*
em_esthum:*
em_estin:*
em_estmu:*
em_estov:*
em_estpl:*
em_estro:*
em_htc:*
gb_est1:*
gb_est2:*
gb_htc:*
gb_est3:*
gb_est4:*
gb_est5:*
em_estfun:*
em_estcom:*
em_gss_hum:*
em_gss_inv:*
em_gss_pln:*
em_gss_vrt:*
em_gss_fun:*
em_gss_mam:*
em_gss_mus:*
em_gss_pro:*
em_gss_rod:*
em_gss_phg:*
em_gss_vrl:*
gb_gss1:*
gb_gss2:*

the number of results predicted by chance to have a
than or equal to the score of the result being printed,
d by analysis of the total score distribution.

SUMMARIES

ry
ch Length DB ID   Description
-- ------ -- --------   -----------
:1   776 13 BX090012   BX090012 BX090012
:8   777 12 BI819200   BI819200 60303464614
:6 1071 12 BM921213   BM921213 AGENCOURT
:1   963 13 BQ671259   BQ671259 AGENCOURT

#	c	ID	n1	n2	pct	n3	ID (suffix)
5	c	CA396679	14	567	41.3	567	CA396679 C
6	c	CB141389	14	545	39.7	545	CB141389 K
7		BI596681	12	828	37.7	518	BI596681 6
8		BI870393	12	728	36.9	506	BI870393 6
9		BI871711	12	731	36.9	506	BI871711 6
10	c	CA413067	14	575	36.4	500	CA413067 U
11	c	BI966060	12	609	35.8	491	BI966060 U
12		CF126539	14	666	33.9	465	CF126539 U
13	c	BU951915	14	474	33.6	462	BU951915 i
14	c	CB529199	13	569	32.7	449	CB529199 U
15		BQ707185	13	948	31.8	437	BQ707185 A
16		BQ884231	13	940	31.8	436	BQ884231 A
17		CF126932	14	697	31.0	425	CF126932 U
18		BI766766	12	834	30.3	416	BI766766 6
19	c	AI422796	8	413	28.7	394	AI422796 tf
20	c	BM971606	12	568	28.7	393	BM971606 U
21	c	BU631264	13	569	28.6	393	BU631264 U
22		BI824443	6	531	27.3	375	BI824443 6
23		CB998034	14	824	26.4	362	CB998034 A
24		BM128059	12	440	25.6	352	BM128059 i
25		BQ674188	13	951	25.4	349	BQ674188 A
26		BM509016	11	377	24.8	341	BM509016 i
27	c	BE858778	7	710	24.0	329	BE858778 7
28	c	BM662622	10	367	23.8	328	BM662622 U
29	c	BE858822	7	367	23.8	327	BE858822 7
30	c	AW195034	10	337	23.7	326	AW195034 x
31	c	BF439993	10	345	23.7	326	BF439993 n
32	c	AW204512	10	352	23.7	326	AW204512 U
33	c	AW291620	10	364	23.7	324	AW291620 U
34	c	AI913541	9	399	23.6	324	AI913541 wa
35		BI966255	12	456	23.6	324	BI966255 i
36		AQ100365	28	367	23.4	321	AQ100365 H
37		BM925491	12	1819	23.4	321	BM925491 A
38		BM505649	12	374	23.3	320	BM505649 i
39	c	AI695776	9	329	23.2	318	AI695776 wb
40	c	BF195436	7	340	23.0	318	BF195436 7
41	c	AI221985	9	407	23.0	316	AI221985 gg
42	c	AI291866	9	416	22.5	309	AI291866 qm
43		BG110063	6	910	22.1	307	BG110063 6
44	c	AI202121	9	317	22.1	303	AI202121 qi
45	c	BM703512	12	346	21.7	298	BM703512 U
46		BU729427	13	333	21.6	296	BU729427 U
47	c	BM688946	12	298	21.1	290	BM688946 U
48		BG686319	12	587	20.6	283	BG686319 6
49		AF163779	28	1027	19.7	271	AF163779 A
50	c	AI760777	9	346	19.4	267	AI760777 wi
51	c	BF940141	10	346	19.1	262	BF940141 n
52	c	BG054878	10	561	18.0	247	BG054878 n
53	c	AW081731	9	318	17.2	236	AW081731 xb
54	c	AW661741	10	264	16.8	230	AW661741 h
55	c	AI863563	9	253	16.7	229	AI863563 wj
56	c	AI682487	9	238	16.5	227	AI682487 wc
57		BI906850	6	698	15.8	217	BI906850 6
58	c	BG054914	10	351	15.4	211	BG054914 n
59	c	AI669243	9	397	15.1	207	AI669243 wc
60	c	BI908274	12	894	15.0	206	BI908274 6
61		AI091441	6	465	14.7	202	AI091441 owf
62	c	BF222608	10	224	13.9	191	BF222608 7
63		C00994	13	282	13.3	182	C00994 HUM
64	c	CF994566	14	1064	12.7	175	CF994566 A
65	c	CF594233	14	1319	12.7	175	CF594233 A
66	c	AA913913	9	338	12.7	174	AA913913 ol
67	c	BE501197	12	196	12.3	169	BE501197 7
68	c	BI677255	12	422	12.2	167	BI677255 i
69	c	BI677256	12	372	11.8	162	BI677256 i
70	c	AQ890280	28	436	11.8	162	AQ890280 H
71	c	AW131279	10	409	11.5	158	AW131279 x
72		R55379	14	345	10.2	140	R55379 yj7
73	c	BI762908	12	785	9.6	132	BI762908 6
74	c	AI865482	9	179	9.5	130	AI865482 wk
75	c	AI081661	9	210	9.3	128	AI081661 ouf
76	c	N35070	14	468	9.3	128	N35070 yy1
77		BI965174	12	474	8.7	119	BI965174 ic

strand cDNA was primed with a Not I - oligo(dT) I
TGTTACCAATCTGAAGTGGAGGGAGCGGCGGCCCCTTTTTTTTTTTTTTT
double-stranded cDNA was ligated to Eco RI adapto
(Pharmacia), digested with Not I and cloned into
and Eco RI sites of a modified pT7T3 vector (Phar
Library went through one round of normalization t
230.Library constructed by Bento Soares and M.Fe
Bonaldo."

ORIGIN

Query Match   52.1%; Score 716; DB 13; Length 776;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; C

Qy   600 TCAGGTGCACTTTGATGAGGGAGGCTGTGTCTACCTGAAGCTGGACTTGCTGGTGC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TCAGGTGCACTTTGATGAGGGAGGCTGTGTCTACCTGAAGCTGGACTTGCTGGTGC

Qy   660 TGTGCTGGCCCTGCGCTGCCTGGAGAATTCTCAGCCACTGCGCGCCAGTTCCCTCC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TGTGCTGGCCCTGCGCTGCCTGGAGAATTCTCAGCCACTGCGCGCCAGTTCCCTCC

Qy   720 CCAGCTCCGCCCTCTGCCAGGTGTCTGGGCTGTTGGCCCTGCGGCCAGGGTCCTCC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CCAGCTCCGCCCTCTGCCAGGTGTCTGGGCTGTTGGCCCTGCGGCCAGGGTCCTCC

Qy   780 GATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCTTCCTCACCTACTTCC
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCTTCCTCACCTACTTCC

Qy   840 CTTCCAGTTCACTGAGGGCCCTGGTCTCCCCACAGTCGTCCCAGGCTGCCGGCT
                |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CTTCCAGTTCACTGAGGGCCCTGGTCTCCCCACAGTCGTCCCAGGCTGCCGGCT

Qy   900 TCGACAGCTCTCTGGGCACCCGGTCCCGTCCCCGCCCACCCTCAGCCGCTCTTTGCT
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TCGACAGCTCTCTGGGCACCCGGTCCCGTCCCCGCCCACCCTCAGCCGCTCTTTGCT

Qy   960 ACCTGCCCCTCCCTCAGAGCTGCCTGGCCTGTTCACGTGTTTCCATCCCACA
                ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 ACCTGCCCCTCCCTCAGAGCTGCCTGGCCTGTTCACGTGTTTCCATCCCACA

Qy   1020 TACAGTATTCCCACTCTTATCTTACACTCCCCACCGCCACTCTCCACTTCACT
                ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TACAGTATTCCCACTCTTATCTTACACTCCCCACCGCCACTCTCCACTTCACT

Qy   1080 CCCCAATCCCTGACCCTTGAGGCCCCAGTGATCTCGACTCCCCCTGGCCACAG
                |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CCCCAATCCCTGACCCTTGAGGCCCCAGTGATCTCGACTCCCCCTGGCCACAG

Qy   1140 CCAGGGCATTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCCCAC
                ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CCAGGGCATTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCCCAC

Qy   1200 GGCACTAAGAGAGGGCTGGACCTGGCCGGCAGGAAGCCAAAGAGACTGGGCCTAGGCC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GGCACTAAGAGAGGGCTGGACCTGGCCGGCAGGAAGCCAAAGAGACTGGGCCTAGGCC

Qy   1260 GTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCTCCCCTTGAGAATTCCCTG
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCTCCCCTTGAGAATTCCCTG

RESULT 2
BI819200
BI819200
LOCUS   BI819200   777 bp   mRNA   linear   EST 04
DEFINITION   603034614F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5175
              mRNA sequence.
ACCESSION   BI819200
VERSION   BI819200.1 GI:15930750
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens

8.7   639   10   BB642326   BB642326 BB642326
8.7   692   13   BY748962   BY748962 BY748962
8.7 1187   12   BQ053284   BQ053284 AGENCOURT
8.7 2237   11   AK044387   AK044387 Mus muscu
7.4 1926   12   BM906056   BM906056 AGENCOURT
7.3   289   10   BE503779   BE503779 AGENCOURT
7.1   254   12   BM128317   BM128317 if12d09.x
6.9   837   14   CB961075   CB961075 AGENCOURT
6.8   105   10   AW191845   AW191845 x176b04.x
6.3   538   10   BF821434   BF821434 MR1-RT003
6.1   140   9   AI582027   AI582027 ar98c06.x
5.9   576   9   AI535912   AI535912 vicpro1.H
5.3   278   10   AW417023   AW417023 52923 MAR
5.3   404   10   BF044430   BF044430 BP250013A
5.3   542   10   BF041509   BF041509 BP250025B
5.2   257   14   R55285   R55285   yj77a08.s1
5.2   784   12   BG686384   BG686384 602638308
4.2   571   10   BF073881   BF073881 220949 MA
3.8   584   10   AW917574   AW917574 EST348878
3.6   426   10   AW236926   AW236926 xm51a10.x
3.6   561   10   AW763237   AW763237 ur70d09.y
3.4   237   29   CG616340   CG616340 OST308362
3.4   269   29   CG564287   CG564287 OST188290

ALIGNMENTS

12 Soares breast 2NbHBst Homo sapiens cDNA clone
       776 bp   mRNA   linear   EST 23-JAN-2003
12 Soares breast 2NbHBst Homo sapiens cDNA clone
98E15243 ; IMAGE:154742, mRNA sequence.
12
12.1 GI:27821952

apiens (human)
apiens
ota;Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 776)
L., Hei,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
f.U., Schneider,D. and Korn,B.
UnigeneSet - RZPD3
ished (2003)
t: Ina Rolfs
eutsches Ressourcenzentrum fuer Genomforschung GmbH
enheimer Feld 580, D-69120 Heidelberg, Germany
IMAGp998E15243.
B; I.M.A.G.E. cDNA Clone Collection;
UnigeneSet - RZPD3 (RZPDLIB No.972)
/www.rzpd.de/CloneCards/cgi-
owLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
eutsches Ressourcenzentrum fuer Genomforschung GmbH
rweg 6, D-14059 Berlin, Germany
49 30 32639 101
49 30 32639 111
pd.de
lone is available royalty-free from RZPD;
t RZPD (clone@rzpd.de) for further information. Seq primer:
Primer sequence: TTTCACACAGGAAACAGCTATGAC.
Location/Qualifiers
1..776
       /organism="Homo sapiens"
       /mol_type="mRNA"
       /db_xref="taxon:9606"
       /clone="IMAGp998E15243 ; IMAGE:154742"
       /sex="Female"
       /dev_stage="adult"
       /lab_host="DH10B (ampicillin resistant)"
       /clone_lib="Soares breast 2NbHBst"
       /note="Organ: breast; Vector: pT7T3D (Pharmacia) with a
       modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st

ca; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
as 1 to 777)
http://mgc.nci.nih.gov/.
l Institutes of Health, Mammalian Gene Collection (MGC)
shed (1999)
: Robert Strausberg, Ph.D.
cgapbs-r@mail.nih.gov
Procurement: Life Technologies, Inc.
ibrary Preparation: Life Technologies, Inc.
ibrary Arrayed by: The I.M.A.G.E. Consortium (LLNL)
quencing by: Incyte Genomics, Inc.
distribution: MGC clone distribution information can be
hrough the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov
LLAM1437 row: 1 column: 03
ality sequence stop: 759.
Location/Qualifiers
    1..777
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:5175698"
    /lab_host="DH10B"
    /clone_lib="NIH_MGC 115"
    /note="Organ: pooled brain, lung, testis; Vector:
    pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
    source anonymous pool of 6 male brains, age range 23-27; 1
    male lung, age 27; and 1 male testis, age 69. Library is
    oligo-dT primed and directionally cloned (EcoRV site is
    destroyed upon cloning). Average insert size 1.8 kb,
    insert size range 1-3 kb. Library is normalized and
    enriched for full-length clones and was constructed by C.
    Gruber (Invitrogen). Research Genetics tracking code
    021. Note: this is a NIH_MGC Library."

rity   45.8%;   Score 629; DB 12; Length 777;
       99.7%;   Pred. No. 3.1e-309;
nservative   0; Mismatches   2;   Indels   2;   Gaps   0;

AGGAGGAGACCAGGACCCGTCGGAACTGAATCCCCAGACAGAGAAAGCCAGGATCC 347
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGGAGGAGACCAGGACCCGTCGGAACTGAATCCCCAGACAGAGAAAGCCAGGATCC   72

CTTTCCTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAAC 407
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTTTCCTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAAC 132

CTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACGG 467
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACGG 192

AGGCAGGTGTGACGGGACAGTGAGTGGCTCGGGAGGAAGCCAGAATCAACAGCTC 527
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGGCAGGTGTGACGGGACAGTGAGTGGCTCGGGAGGAAGCCAGAATCAACAGCTC 252

CTCTGCGCTACAACCGCCAGATCGGGAGTTTATAGTCACCCGGCTGGGCTCTA 587
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTCTGCGCTACAACCGCCAGATCGGGAGTTTATAGTCACCCGGCTGGGCTCTA 312

TGTACTGTCAGGTGCACTTGATGAGGGAAGGCTGTCTACCTGAAGCTGGACTT 647
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGTACTGTCAGGTGCACTTGATGAGGGAAGGCTGTCTACCTGAAGCTGGACTT 372

TGGATGGTGTGTCGGCCCTGCCCTGGCCTGGACGAATTCTCAGCCACTGCGGCCAG 707
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGGATGGTGTGTCGGCCCTGCCCTGGCCTGGACGAATTCTCAGCCACTGCGGCGAG 432

TCGGGGCCCAGCTCGCCCTCTGCCAGGTGTCTGGGCGTTGGCCCTGCGGCCAGG 767
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCGGGGCCCCAGCTCCGCCCTCTGCCAGGTGTCTGGGCGTTGGCCCTGCGGCCAGG 492

Qy   768 GTCCTCCCTGCGGATCGGCACCCTCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCT
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   493 GTCCTCCCTGCGGATCGGCACCCTCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCT

Qy   828 CTACTTCGGACTCTTCCAGGTTCACTGAGGGGCCTGGTCTCCCCACAGTCGTCCCA
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   553 CTACTTCGGACTCTTCCAGGTTCACTGAGGGGCCTGGTCTCCCCGCAGTCGTCCCA

Qy   888 TGCCGGCGCCCCCTGCGACAGCTGTCGGGCACCCGGTCCCCTGCCCCACCCCAGC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   613 TGCCGGCGTCCCCCTGCGACAGCTCTGTCGGGCACCCGGTCCCCTGCCCCACCCCAGC

Qy   948 TCTTTGCTCCAGACCTGCCCCTCCCTCTAGAGGCTGCCTGGGGCTGTTCACGTGTTT
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   673 TCTTTGCTCCAGACCTGCCCCTCCCTCTAGAGGCTGCCTGGGGCTGTTCACGTGTTT

Qy   1008 ATCCCACATAA 1018
          |||||||||||
Db   733 ATCCCACATAA 743

RESULT 3
BM921213
LOCUS   BM921213   1071 bp   mRNA   linear   EST 12-
DEFINITION AGENCOURT_6633046 NIH_MGC_115 Homo sapiens cDNA clone IMAGE
            5', mRNA sequence.
ACCESSION   BM921213
VERSION   BM921213.1 GI:19371592
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleo
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1071)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (M
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12786 row: p column: 02
            High quality sequence stop: 656.
FEATURES   Location/Qualifiers
     source   1..1071
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5752561"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC 115"
                     /note="Organ: pooled brain, lung, testis; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroye
                     source anonymous pool of 6 male brains, age range
                     male lung, age 27; and 1 male testis, age 69. Lib
                     oligo-dT primed and directionally cloned (EcoRV si
                     destroyed upon cloning). Average insert size 1.8
                     insert size range 1-3 kb. Library is normalized an
                     enriched for full-length clones and was construct
                     Gruber (Invitrogen). Research Genetics tracking
                     021. Note: this is a NIH_MGC Library."
ORIGIN
Query Match   44.6%;   Score 613; DB 12; Length 1071;
Best Local Similarity   99.7%; Pred. No. 4.8e-301;
Matches 713; Conservative   0; Mismatches   2; Indels   0; Ga

Qy   412 GCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACGGA

Plate: LLCM2456   row: 1   column: 13
High quality sequence stop: 565.

FEATURES   Location/Qualifiers
     source   1..963
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6274716"
                     /tissue_type="epidermoid carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_102"
                     /note="Organ: salivary gland; Vector: pOTB7; Site
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using
                     following 5' adaptor: GGCACGAG(G). Library const
                     by Ling Hong in the laboratory of Gerald M. Rubin
                     (University of California, Berkeley) using ZAP-c
                     synthesis kit (Stratagene) and Superscript II RT
                     Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match   44.1%;   Score 605;   DB 13;   Length 963;
Best Local Similarity   100.0%;   Pred. No. 5.7e-297;
Matches 605;   Conservative   0;   Mismatches   0;   Indels   0;   G

Qy   408 ACGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGACACAGG
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   39 ACGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGACACAGG

Qy   468 AGCGCAGGCAGGTGTGACGGGACAGTGAGTGGCTCGGGAGGAAGCCAGAATCAACA
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   99 AGCGCAGGCAGGTGTGACGGGACAGTGAGTGGCTCGGGAGGAAGCCAGAATCAACA

Qy   528 CAGCCCTCTGCGCTACAACCGCCAGATCGGGAGTTTATAGTCACCGGGCTGGGC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   159 CAGCCCTCTGCGCTACAACCGCCAGATCGGGAGTTTATAGTCACCGGGCTGGGC

Qy   588 CTACTTGTACTGTCAGGTGCACTTTGATGAGGGGAGGGCTGTCTACCTGAAGCTGG
         ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   219 CTACCTGTACTGTCAGGTGCACTTTGATGAGGGGAGGGCTGTCTACCTGAAGCTGG

Qy   648 GCTGGTGGATGGTGTCGTGCTGGCCTGCGCTGCCTGAGGAATTCTCAGCCACTGCGG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   279 GCTGGTGGATGGTGTCGTGCTGGCCTGCGCTGCCTGAGGAATTCTCAGCCACTGCGG

Qy   708 TTCCCTCGGGGCCCCAGCTCCGCTCCGCTCCCCAGGTGTCTCGGGCGTGTTGGCCCTCGCGGC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   339 TTCCCTCGGGGCCCCAGCTCCGCTCCGCTCCCCAGGTGTCTCGGGCGTGTTGGCCCTCGCGGC

Qy   768 GTCCTCCCTGCGGATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   399 GTCCTCCCTGCGGATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCC

Qy   828 CTACTTCGGACTCTTCCAGGTTCACTGAGGGGCCCTGGTCTCCCCACAGTGTCCC
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   459 CTACTTCGGACTCTTCCAGGTTCACTGAGGGGCCCTGGTCTCCCCACAGTGTCCC

Qy   888 TGCCGGCTCCCCTGACAGCTCTCTGGGCACCCGGTCCCCCTGCCCACCCCTCAG
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   519 TGCCGGCTCCCCTGACAGCTCTCTGGGCACCCGGTCCCCCTGCCCACCCCTCAG

Qy   948 TCTTTGCTCCAGACCTGCCCCTCCCCTCCTAGAGGCCTGGGCCTGTTCACGTGTT
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   579 TCTTTGCTCCAGACCTGCCCCTCCCCTCCTAGAGGCCTGGGCCTGTTCACGTGTT

Qy 1008 ATCCC 1012
         |||||
Db   639 ATCCC 643

RESULT 5
CA396679
CA396679
LOCUS   CA396679   567 bp   mRNA   linear   EST 06

Qy   GAAGAGCGGATCCAGCCCATTATGAAGTTCATCCACGACCTGGACACGACCGGAGCG 71
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db

Qy   CAGGTGTGGACGGGACAGTGAGTGGCTCGGGAGGAAGCCAGAATCAACAGCTCCAGC 531
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   CAGGTGTGGACGGGACAGTGAGTGGCTCGGGAGGAAGCCAGAATCAACAGCTCCAGC 131

Qy   TGCGCTACAACCGCCAGATCGGGAGTTTATAGTCACCGGGCTGGCTCTACTAC 591
      |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   TGCGCTACAACCGCCAGATCGGGAGTTTATAGTCACCGGGCTGGCTCTACTAC 191

Qy   ACTGTCAGGTGCACTTTGATGAGGGGAAGGCGTGTCTACCTGAAGCTGGACTTGCTG 651
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   ACTGTCAGGTGCACTTTGATGAGGGGAAGGCGTGTCTACCTGAAGCTGGACTTGCTG 251

Qy   ATGGTGTGCTGGCCCTGCGCTGCCTGAGGAATTCTCAGCCACTGCGGCCAGTTCC 711
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   ATGGTGTGCTGGCCCTGCGCTGCCTGAGGAATTCTCAGCCACTGCGGCGAGTTCC 311

Qy   GGCCCAGCTCCGCTCCGCTCCCCTGGGCTGTCTGGGCCCTGCGGCCAGGGTCC 771
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   GGCCCAGCTCCGCTCCGCTCCCCTGGGCGTGTCTGGGCCCTGCGGCCAGGGTCC 371

Qy   TGCGGATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACCTAC 831
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   TGCGGATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACCTAC 431

Qy   GACTCTTCCAGGTTCACTGAGGGGCCCTGGTCTCCCCACAGTGTCTCCCAGGCTGCC 891
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   GACTCTTCCAGGTTCACTGAGGGGCCCTGGTCTCCCCCAGTGTCGTCCCAGGCTGCC 491

Qy   CCCCTGACAGCTCTCTGGGCACCCGGTCCCCCTCGCCCCACCCCTCAGCCGGCTCTT 951
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   CCCCTGACAGCTCTCTGGGCACCCGGTCCCCCTCGCCCCACCCCTCAGCCGCTCTT 551

Qy   CCAGACCTGCCCCTCCCTCTAGAGGCTGCGCTGGGCTGTTCCACGTGTTTCCATCC 1011
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   CCAGACCTGCCCCTCCCTCTAGAGGCTGCGCTGGGCTGTGTTTCACGTGTTTCCATCC 611

Qy   TAAATACAGTATTCCCACTCTTATCTTACACAACTCCCCACCTCTCCCACCT 1071
      ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   TAAATACAGTATTCCCACTCTTATCTTACACAACTCCCCACCTCTCCCACCT 671

Qy   AGCTCCCCAATCCCTGACCCTTTGAGGCCCCAGTGATCTCGACTCCCCCC 1126
      ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   AGCTCCCCAATCCCTGACCCTTTGAGGCCCCAGTGATCTCGACTCCCCCC 726

59   963 bp   mRNA   linear   EST 15-JUL-2002
JRT 8303564 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6274716
NA sequence.
59
59.1 GI:21782093

apiens (human)
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
C http://mgc.nci.nih.gov/.
al Institutes of Health, Mammalian Gene Collection (MGC)
ished (1999)
t: Robert Strausberg, Ph.D.
cgapbs-r@mail.nih.gov
Procurement: ATCC
Library Preparation: Rubin Laboratory
Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
equencing by: Agencourt Bioscience Corporation
distribution by: MGC clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL at:
/image.llnl.gov

```
Qy    295  GAGGACCAGGACCCGTCGGAACTGAATCCCCAGACAGAAGAAAGCCAGGATCCTGCC
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  GAGGACCAGGACCCGTCGGAACTGAATCCCCAGACAGAAGAAAGCCAGGATCCTGCC

Qy    355  TTCCTGACCGACTAGTTCGGCCTCGCGAGAAGTGCACCTAAAGGCCGAAAAACACGC
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  TTCCTGACCGACTAGTTCGGCCTCGCGAGAAGTGCACCTAAAGGCCGAAAAACACGC

Qy    415  CGAAGAGGCGATCGCAGCCCATTATGAAGTCATTCACGACCTGGACAGGACGGAGCC
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  CGAAGAGGCGATCGCAGCCCATTATGAAGTCATTCACGACCTGGACAGGACGGAGCC

Qy    475  GCAGGTGTGGACGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAGC
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  GCAGGTGTGGACGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAGC

Qy    535  CTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGCTGGGCTCTACTAC
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  CTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGCTGGGCTCTACTAC

Qy    595  TACTGTCAGGTGCACTTTGATGAGGGG  621
               |||||||||||||||||||||||||||
Db    541  TACTGTCAGGTGCACTTTGATGAGGGG  567
```

RESULT 6
CB141389
LOCUS   CB141389   545 bp   mRNA   linear   EST 29-
DEFINITION K-EST0194999 L15CKK1 Homo sapiens cDNA clone L15CKK1-30-E06
           mRNA sequence.
ACCESSION   CB141389
VERSION   CB141389.1 GI:28116436
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelec
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 545)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., K
          Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S
          Kim.Y.S.
  TITLE   21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT   Contact: Kim YS
          Genome Research Center
          Korea Research Institute of Bioscience & Biotechnology
          52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
          Tel: +82-42-860-4470
          Fax: +82-42-860-4409
          Email: yongsung@mail.kribb.re.kr
          Plate: 30 row: E column: 06
          High quality sequence stop: 545.
FEATURES   Location/Qualifiers
    source   1..545
          /organism="Homo sapiens"
          /mol_type="mRNA"
          /db_xref="taxon:9606"
          /clone="L15CKK1-30-E06"
          /sex="M"
          /cell_line="CK-K1"
          /lab_host="Top10F'"
          /clone_lib="L15CKK1"
          /note="Organ: Liver; Vector: pCNS-D2; Site 1: EcoR
          Site 2: NotI; The poly (A)+ RNA was dephosphorylat
          bacterial alkaline phosphatase (BAP) and then deca
          with tabacco acid pyrophosphatase (TAP). The decap
          intact mRNA was ligated with DNA-RNA linker includ
          EcoRI site by treatment of T4 RNA ligase and the f
          strand cDNA was synthesized from oligo dT-selected
          priming with dT-tailed vector. The dT-tailed vecto
          adjusted to have about 60nt. The cDNA vector was
          circularized with E. coli DNA ligase after digesti

.y1 Human Retinal pigment epithelium/choroid cDNA
malized, unamplified) : cs Homo sapiens cDNA clone cs80h07
A sequence.
9
9.1 GI:24733297

piens (human)
piens
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
G., Bernstein,S.L., Wyatt,M.K., Farris,R.N., Behal,A.,
n,J.W., Boffard,G., Smith,D. and Peterson,K.
ed sequence tag analysis of human RPE/choroid for the
Project: Over 6000 non-redundant transcripts, novel genes
ice variants
s. 8 (4), 205-220 (2002)
0
0
: Wistow G
on Molecular Structure and Function
1 Eye Institute
NIH, Bethesda, MD 20892-2740, USA
1 402 3452
1 496 0078
graeme@helix.nih.gov
80 row: h column: 07
mer: M13RP1 reverse primer (ABI).
    Location/Qualifiers
    1..567
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="cs80h07"
    /tissue_type="RPE/choroid"
    /dev_stage="Adult"
    /lab_host="EMDH10B"
    /clone_lib="Human Retinal pigment epithelium/choroid cDNA
    (Un-normalized, unamplified): cs"
    /note="Organ: Eye; Vector: pCMVSPORT6; Two different donor
    eyes (75-80 years old) yielded approximately 600 mg of
    dissected RPE/choroid tissue. This in turn yielded 340 ug
    of total RNA and 7 ug of mRNA. A directionally cloned cDNA
    library in the pCMVSPORT6 vector was constructed at Life
    Technologies (Rockville, MD; now part of Invitrogen Corp),
    essentially following the protocols of the SuperScript
    Plasmid System (Invitrogen Corp.
    <http://www.invitrogen.com/>). The library code
    designation was cs. For this library, cDNA inserts were
    cloned into the NotI/MluI sites of the vector. EST
    analysis was performed on the unamplified library at the
    NIH Intramural Sequencing Center (NISC)."

rity   41.3%; Score 567; DB 14; Length 567;
nservative   100.0%; Pred. No. 1.3e-277;
               0; Mismatches   0; Indels   0; Gaps   0;
```

```
GTCCCGGGATGGGGGGCGGGTGAGGCAGGCACAGCCCCCCGCCCCCATGGCCGCC 114
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTCCCGGGATGGGGGGCGGGTGAGGCAGGCACAGCCCCCCGCCCCCATGGCCGCC 60

GAGGCCAGGACGGCGAGGGGGCCGCCGGGGGAGCCGGGGGAGCCCCGGGTCCCG 174
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GAGGCCAGGACGGCGAGGGGGCGCCGGGGGGAGCCGGGGGAGCCGGGGTCCCG 120

GCTCGGGGCCTGGCGCCTGGCGCTCGCCTCCGGCCTCCGCTGCTGCCCGGTCAGT 234
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCTCGGGGCCTGGCGCCTGGCGCTCGCCTCCGGCCTCCGCTGCTGCCCGGTCAGT 180

GAGGCCGGGCATCGCTGTTCCGCCCAGGAGCTCGTCCCAGGAGGAGCTCGTGGCCAGAG 294
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GAGGCCGGGCATCGCTGTTCCGCCCAGGAGCCTCGTCCCAGGAGGAGCTCGTGGCAGAG 240
```

EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

```
arity      39.7%;   Score 545;   DB 14;   Length 545;
           100.0%;  Pred. No. 2.1e-266;
onservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
Qy   GGAACTGAATCCCCAGACAGAAGAAAGCCAGGATCCTCGCGCCTTTCCTGAACCGACT  368
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   GGAACTGAATCCCCAGACAGAAGAAAGCCAGGATCCTCGCGCCTTTCCTGAACCGACT   60

Qy   CGGCTCCTCGCAGAAGTGCACCTAAAGGCCGGAAAACACGGGCTCGAAGAGCGATCGC  428
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   CGGCTCCTCGCAGAAGTGCACCTAAAGGCCGGAAAACACGGGCTCGAAGAGCGATCGC  120

Qy   CATTATGAAGTTCATCCACGACCTGGACAGGACGGAGCGCAGGCAGGTGTGGACGG  488
     |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   CATTATGAAGTTCATCCACGACCTGGACAGGACGGAGCGCAGGCAGGTGTGGACGG  180

Qy   GTGAGTGGCTGGGAGGAGAAGCCAGAATCAACAGTCCAGCCCCTCGCCTACAACCG  548
     |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   GTGAGTGGCTGGGAGGAGAAGCCAGAATCAACAGTCCAGCCCCTCGCCTACAACCG  240

Qy   ATCGGGGAGTTATAGTCACCCGGGCTGGGCTCTACTACCTGTACTGGCTCAGGTGCA  608
     |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   ATCGGGGAGTTATAGTCACCCGGGCTGGGCTCTACTACCTGTACTGGCTCAGGTGCA  300

Qy   GATGAGGGGAAGGCTGTCTCACCTGAAGCTGGACTTGGTGGTGTGCTGGC  668
     |||||||||||||||||||||||||||||||||||||||||||||||||
Db   GATGAGGGGAAGGCTGTCTCACCTGAAGCTGGACTTGGTGGTGTGCTGGC  360

Qy   CGCTGCCTGGAGGAATTCTCAGCCACTGCGGCCACTGCGGCCAGTTCCCTCGGCCCCAGTCCCG  728
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   CGCTGCCTGGAGGAATTCTCAGCCACTGCGGCCACTGCGGCCAGTTCCCTCGGCCCCAGTCCCG  420

Qy   TGCCCAGGTGTCTCGGGGTGTTGGCCCTGCGGCCCGCCAGGGTCCTCCCTGCGGATCCGCAC  788
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   TGCCCAGGTGTCTCGGGGTGTTGGCCCTGCGGCCCGCCAGGGTCCTCCCTGCGGATCCGCAC  480

Qy   CCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACCTACTTCCGGACTCTTCCAGGT  848
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   CCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACCTACTTCCGGACTCTTCCAGGT  540

Qy   T  853
   
Db   T  545
```

81
254Fl NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5285892 5',
sequence.
81
81.1   GI:15489620

apiens (human)
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
uses 1 to 828)
C http://mgc.nci.nih.gov/.
al Institutes of Health, Mammalian Gene Collection (MGC)
ished (1999)

t: Robert Strausberg, Ph.D.
cgapbs-r@mail.nih.gov
; Procurement: Miklos Palkovits, M.D., Ph.D.
Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information ca
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11722 row: k column: 13
High quality sequence stop:776.

```
FEATURES             Location/Qualifiers
     source          1..828
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5285892"
                     /tissue_type="hypothalamus"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_96"
                     /note="Organ: brain; Vector: pBluescriptR (modifi
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-Xho
                     (gtcgag); Oligo-dT primed using primer
                     5'-TTTTTTTTTTTTTTVN-3', size-selected for  aver
                     insert size 2.3 kb and normalized to ROT 5...This
                     primary library enriched for full-length clones
                     constructed using the Cap-trapper method (Carninc
                     preparation).  Library constructed by M. Brownste
                     (NIMH/NHGRI, National Institutes of Health). Note
                     a NIH_MGC Library."
```

```
ORIGIN

Query Match           37.7%;    Score 518;  DB 12;   Length 828;
Best Local Similarity 99.7%;    Pred. No. 1.4e-252;
Matches 618; Conservative  0;  Mismatches  2;  Indels  0;  G

Qy   312  GGAACTGAATCCCCAGACAGAGAGAAAGCCAGGATCCTCGCGCCTTTCCTGAACCGAC
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    43  GGAACTGAATCCCCAGACAGAGAGAAAGCCAGGATCCTCGCGCCTTTCCTGAACCGAC

Qy   372  TCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAACACGGGCTCGAAGAGCGATCC
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   103  TCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAACACGGGCTCGAAGAGCGATCC

Qy   432  CCATTATGAAGTTCATCCACGACCTGGACAGGACGGAGCGCAGGCAGGTGTGGACC
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   163  CCATTATGAAGTTCATCCACGACCTGGACAGGACGGAGCGCAGGCAGGTGTGGACC

Qy   492  AGTGAGTGGCTGGGAGGAGAAGCCAGAATCAACAGTCCAGCCCCTCGCCTACAACC
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   223  ATTGAGTGGCTGGGAGGAGAAGCCAGAATCAACAGTCCAGCCCCTCGCCTACAAC

Qy   552  GATCGGGGGAGTTATAGTCACCCGGGCTCTACTACCTGTACTGTCAGGTGC
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   283  GATCGGGGGAGTTATAGTCACCCGGGCTCTACTACCTGTACTGTCAGGTGC

Qy   612  TGATGAGGGGAAGGCTGTCTCACCTGAACTGGACTTGGCTGGATGGTGTGCTGC
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   343  TGATGAGGGGAAGGCTGTCTCACCTGAACTGGACTTGGCTGGATGGTGTGCTGC

Qy   672  GCGCTGCCTGGAGGAATTCTCAGCCACTGCGGCCACTGCGGCCCCCAGTCC
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   403  GCGCTGCCTGGAGGAATTCTCAGCCACTGCGGCCACTGCGGCCCCCAGTCC

Qy   732  CTGCCAGGTGTCTGGGCTGTTGGGCGTGGTGTGGTCCTGCGCCAGGGTCCTCCCTGCGGATCCGCA
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   463  CTGCCAGGTGTCTGGGCTGTTGGCCTGCGTCCCTGCGGGTCCAGGGTCCTCCCTGCGGATCCGCA

Qy   792  CCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACCTACTTCGGACTCTTCCAGG
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   523  CCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACCTACTTCGGACTCTTCCAGG

Qy   852  CTGAGGGGCCCTGGTCTCCCCACAGTCGTCCCAGGCTGCCGGCTCCCCTGCGCAGC
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   583  CTGAGGGGCCCTGGTCTCCCCAGTCGTCCCAGGCTGCCGGCTCCCCTGCGCAGC
```

```
Db    931  ACCCGGTCCCCTCTG
           ||||||||||||||
Qy    662  ACCCGGTCCCCTCTG

           3       728 bp    mRNA     linear  EST 11-OCT-2001
           41F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5405459 5',
           quence.
           3
           3.1  GI:16044066

           piens (human)
           piens
           ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
           es 1 to 728)
           'http://mgc.nci.nih.gov/.
           l Institutes of Health, Mammalian Gene Collection (MGC)
           shed (1999)
           : Robert Strausberg, Ph.D.
           cgapbs-r@mail.nih.gov
           Procurement: ATCC
           ibrary Preparation: Life Technologies, Inc.
           ibrary Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           quencing by: Incyte Genomics, Inc.
           distribution: MGC clone distribution information can be
           hrough the I.M.A.G.E. Consortium/LLNL at:
           image.llnl.gov
           LLAM12034  row: a  column: 12
           ality sequence stop:728.
           Location/Qualifiers
           1..728
                 /organism="Homo sapiens"
                 /mol_type="mRNA"
                 /db_xref="taxon:9606"
                 /clone="IMAGE:5405459"
                 /tissue_type="adenocarcinoma, cell line"
                 /lab_host="DH10B (phage-resistant)"
                 /clone_lib="NIH_MGC_90"
                 /note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
           Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
           Average insert size 1.7 kb. Library enriched for
           full-length clones and constructed by Life Technologies.
           Note: this is a NIH_MGC Library."

           rity   36.9%;   Score 506;   DB 12;  Length 728;
           100.0%;  Pred. No. 1.8e-246;  Indels   0;  Gaps     0;
           nservative   0;  Mismatches   0;  Indels   0;  Gaps     0;

Qy    261  GCCTGCCCAGGAGGAGCTGGTGCAGAGGAGGACCAGGACCCGTCGGAACTGAAT  321
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         GCCTGCCCAGGAGGAGCTGGTGCAGAGGAGGACCAGGACCCGTCGGAACTGAAT  133

Qy         GACAGAGAAAGCCAGGATCCTGCGCCTTTCCTGAACCGACTAGTTCGGCCTCGC  381
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         GACAGAGAAAGCCAGGATCCTGCGCCTTTCCTGAACCGACTAGTTCGGCCTCGC  193

Qy         TGCACCTAAAGGCCGGAAAACACGGGCTCGAAGAGCGATCGCAGCCCATTATGAA  441
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         TGCACCTAAAGGCCGGAAAACACGGGCTCGAAGAGCGATCGCAGCCCATTATGAA  253

Qy         TCCACGACCTGGACAGGACGGAGCCGCAGGCAGTGTGGACGGGACAGTGAGTGGC  501
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         TCCACGACCTGGACAGGACGGAGCCGCAGGCAGTGTGGACGGGACAGTGAGTGGC  313

Qy         GGAAGCCAGAATCAACAGCTCCAGCCCTCGCGCCCTCTGCGCTACAACCGGCAGATCGGGGAG  561
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         GGAAGCCAGAATCAACAGCTCCAGCCCTCGCGCCCTCTGCGCTACAACCGGCAGATCGGGGAG  373

Qy         AGTCACCCGGGCTGGGCTTCTACTACCTGTACTGTCAGGTGCACTTTGATGAGGGG  621
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    374  TTTATAGTCACCCGGGCTGGGCTTCTACTACCTGTACTGTCAGGTGCACTTGATGA(
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    622  AAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGGATGGTGTGCTGGCCCTGCGCTG(
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    434  AAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGGATGGTGTGCTGGCCCTGCGCTG(

Qy    682  GAGGAATTCAGCCACTGCGGCCAGTTCCTCGGGGCCCAGTTCCCGCCTCTGCCA(
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    494  GAGGAATTCAGCCACTGCGGCCAGTTCCTCGGGGCCCCAGTTCCCGCCTCTGCCA(

Qy    742  TCTGGGCTGTTGGCCCTGCGGCCAGG  767
           ||||||||||||||||||||||||||
Db    554  TCTGGGCTGTTGGCCCTGCGGCCAGG  579
```

```
RESULT 9
BI871711
BI871711
LOCUS        BI871711           731 bp    mRNA     linear  EST 11-
DEFINITION   603395825F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:54054:
             mRNA sequence.
ACCESSION    BI871711
VERSION      BI871711.1   GI:16045386
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteled
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 731)
  AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
  TITLE      National Institutes of Health, Mammalian Gene Collection (b
  JOURNAL    Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLAM12034  row: b  column: 07
             High quality sequence stop: 728.
FEATURES             Location/Qualifiers
     source          1..731
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5405478"
                     /tissue_type="adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_90"

ORIGIN
           Query Match      36.9%;   Score 506;  DB 12;  Length 731;
           Best Local Similarity   100.0%;  Pred. No. 1.8e-246;
           Matches   506;  Conservative   0;  Mismatches   0;  Indels   0;  Ga

Qy    262  CAGGAGCCTGCCCAGGAGGAGCTGGTGCAGAGGAGGACCAGGACCCGTCGGAACTG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     74  CAGGAGCCTGCCCAGGAGGAGCTGGTGCAGAGGAGGACCAGGACCCGTCGGAACTG

Qy    322  CCCCAGACAGAAGAAAGCCAGGATCCTGCGCCTTTCCTGAACCGACTAGTTCGGCCT
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    134  CCCCAGACAGAAGAAAGCCAGGATCCTGCGCCTTTCCTGAACCGACTAGTTCGGCCT

Qy    382  AGAAGTGCACCTAAAGGCCGGAAAACACGGGCTCGAAGAGCGATCGCAGCCCATTA1
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
                                           synthesis was primed with an oligo-dT primer cont
                                           Not I site. Double stranded cDNA was ligated to a
                                           adaptor, digested with Not I, and cloned directio
                                           into pT7T3-Pac vector. The oligonucleotide used t
                                           the synthesis of first-strand cDNA contains a lib
                                           sequence that is located between the Not I site a
                                           (dT)18 tail. The sequence tag for this library is
                                           TGATCACGCT.
                                           TAG_TISSUE=grade-2-chondrosarcoma
                                           TAG_LIB=UI-H-EZO
                                           TAG_SEQ=ATCTAATATG"

ORIGIN

        Query Match
        Best Local Similarity   36.4%;   Score 500;   DB 14;   Length 575;
        Matches   500;   Conservative   0;   Mismatches   0;   Indels   0;   C
                   100.0%;   Pred. No. 2e-243;

Qy   874  CAGTCGTCCCAGCTGCGCGCTCCCCTCGAGACGTCTCTGGGGACCCGGTCCCCTC
          |||||||||||||||||||||||||||||| |||||| |||| ||||||||||||
Db   529  CAGTCGTCCCAGCTGCGCGCTCCCCTCGACAGTCTCTCGGGCACCCGGTCCCCTC

Qy   934  CCACCCTCAGCGCGCTCTTGCTCCAGACCTGCCCCTCCCTTAGAGAGCTGCTGGG
          |||||||||||||||||||| |||||||||||||||||||||||||  ||||||
Db   469  CCACCCTCAGCGCGCTCTTTGCTCCAGACCTGCCCCTCCCTTAGAGGCTGCTGGC

Qy   994  TTCACGTGTTTTCCATCCCACCACATAAAATACAGATTCCCACTCTTATCTTACAACTC
          ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   409  TTCACGTGTTTTCCATCCCACCACATAAATACAGATTCCCACTCTTATCTTACAACTC

Qy  1054  ACCGGCCCACTCTCCACCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCCAG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   349  ACCGGCCCACTCTCCACCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCCAG

Qy  1114  CTCGACTCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   289  CTCGACTCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGG

Qy  1174  GGATGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGGCGGCAG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   229  GGATGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGGCGGCAG

Qy  1234  CCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGA
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   169  CCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGA

Qy  1294  CTCCTCCCTTGAGAATTCCCCTGTGGATTTTTAAAACAGATATTTATTTTTATTATTA
          |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   109  CTCCTCCCTTGAGAATTCCCCTGTGGATTTTTAAAACAGATATTATTTTTATTATTA

Qy  1354  GACAAAATGTTGATAAATGG  1373
          ||||||||||||||||||||
Db    49  GACAAAATGTTGATAAATGG  30
```

```
GTGCACCTAAAGGCCGGAAAACACGGGCTCGAAGAGCGATCGAGCCCATTATGAA  253

Qy  ATCCACGACCTGACACGACGGAGCGCGAGGCGCAGGTGTGACGGGACAGTGAGTGGC  501
    |||||||||||||| |||||||||||||||||||||||  |||||||||||||||||
Db  ATCCACGACCTGACAGGACGGAGCGCGAGGCGCAGGTGTGGACGGGACAGTGAGTGGC  313

Qy  AGGAAGCCAGAATCAACAGTCCAGCCCTTCTGCGCTACAACCGCCAGATCGGGGAG  561
    ||||||||||||||||||||||||||| |||| |||||||||||||||||||||||
Db  AGGAAGCCAGAATCAACAGTCCAGCCCCTCTGGCTACAACCGCCAGATCGGGGAG  373

Qy  TAGTCACCCGGCTGGGCTCCTACTACCTGTACTGTCAGGTGCACTTTGATGAGGGG  621
    |||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db  TAGTCACCCGGCTGGGCTCTCACTACCTGTACTGTCAGGTGCACTTTGATGAGGGG  433

Qy  CTGTCTACCTGAAGCTGACTTGCTGGTGTGTGTCTGGCCCTGCGCTGCCTG  681
    |||||||||||||||||||||||||||||    ||||||||| |||||||||||
Db  CTGTCTACCTGAAGCTGACTTGCTGGTGGATGGTGTGCTGGCCCCTGCGCTGCCTG  493

Qy  AATTCTCAGCCACTGCGGCCAGTTCCCTCGGCCCCAGCTCCGCCCTTGCCAGGTG  741
    ||||||||||||||||||||||||||||||| |||||||||||||| |||||||
Db  AATTCTCAGCCACTGCGGCCAGTTCCCTCGGGCCCCAGCTCCGCCCTCTGCCCAGGTG  553

Qy  GGCTGTTCGGCCCTGCGCGGCCAGG  767
    |||||||||||||||||||||||||
Db  GGCTGTTCGGCCCTGCGCGGCCAGG  579
```

```
67     Z0-bap-i-03-0-UI s1 NCI_CGAP_Ch1 Homo sapiens   linear   EST 07-NOV-2002
       Z0-bap-i-03-0-UI 3', mRNA sequence.
67
67.1   GI:24775718
apiens (human)
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 575)
AP http://www.ncbi.nlm.nih.gov/ncicgap.
al Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Gene Index
ished (1997)
t: Robert Strausberg, Ph.D.
cgapbs-r@mail.nih.gov
Procurement: Dr. Steven Gitelis/ Rush Presbiterian, Dept. of
pedics
Library preparation: Dr. M. Bento Soares, University of Iowa
Library Arrayed by: Dr. M. Bento Soares, University of Iowa
equencing by: Dr. M. Bento Soares, University of Iowa
Distribution: Clone distribution information can be obtained
r. M. Bento Soares, bento-soares@uiowa.edu
ollowing repetitive elements were found in this cDNA
ce: 1-84, >AT_rich#Low_complexity (matched compliment)
imer: M13 FORWARD
Yes.
Location/Qualifiers
1..575
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-EZO-bap-i-03-0-UI"
/tissue_type="Chondrosarcoma Grade II"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_Ch1"
/note="Organ: Left pelvis; Vector: pT7T3-Pac (Pharmacia)
with a modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI_CGAP_Ch1 is a cDNA library containing the following
tissue(s): Chondrosarcoma Grade II. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
```

```
RESULT 11
BI966060/c
LOCUS       BI966060          609 bp      mRNA       linear      EST 12
DEFINITION  ie72g04.x1 Melton Normalized Human Islet 4 N4-HIS 1 Homo s
            cDNA clone IMAGE:5672623 3' similar to TR:O54907 O54907 TN
            WEAK INDUCER OF APOPTOSIS ;, mRNA sequence.
ACCESSION   BI966060
VERSION     BI966060.1   GI:16340465
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 609)
REFERENCE
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestne
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifto
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blista
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvil
```

s,T., Jackson,Y. and Bowers,Y.
he Pancreas Consortium
shed (2000)
: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
he Pancreas Consortium
University. Howard Hughes Medical Institute
Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
8
7-495-1812
7-495-8557
dmelton@bichp.harvard.edu
was constructed by Dr. Douglas Melton DNA sequencing by:
ton University Genome Sequencing Center For information on
ng a clone please contact: Juliana Brown
fas.harvard.edu) This sequence now available from the IMAGE
ium, for clone orders contact: info@image.llnl.gov
ality sequence stop: 412.
Location/Qualifiers
   1..609
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /clone="IMAGE:5672623"
   /sex="Both"
   /tissue_type="Islets of Langerhans"
   /dev_stage="Adult"
   /lab_host="DH10B"
   /clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
   /note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;
   Site 2: Sal 1; Starting library constructed using
   SuperScript Plasmid Library kit (Life Technologies). cDNA
   made by oligo-dT priming. Size-selected by column
   fractionation; average insert size 1.08 kb. Library was
   amplified once on solid support and plasmid DNA from
   library was prepared. The library DNA was normalized by
   method #4 from Bonaldo, Lennon, and Soares 1996 Genome
   Research 6:791-806; 0.5 microgram single-stranded library
   plasmid DNA was mixed with 5 micrograms PCR product
   representing library inserts and hybridized to an Ecot of
   20. Single-stranded (unhybridized) plasmids were isolated
   by hydroxyapatite chromatography and used to make this
   library."

rity   35.8%;   Score 491;   DB 12;   Length 609;
rity 100.0%;   Pred. No. 7.9e-239;   Indels   0;   Gaps   0;
nservative   0;   Mismatches   0;

TGCCGGCTCCCCTGACAGCTCTCTCGGGCACCCGGTCCCCTGCCCCACCCTCA 942
||||||||||||||||||||||||||||||||||||||||||||||||||||
TGCCGGCTCCCCTGACAGCTCTCTCGGGCACCCGGTCCCCTGCCCCACCCTCA 454

TCTTTGCTCCAGACCTGCCCCCTCCCTCTAGAGGCTGCCTGGGCCTGTTCACGTGT 1002
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCTTTGCTCCAGACCTGCCCCCTCCCTCTAGAGGCTGCCTGGGCCTGTTCACGTGT 394

ATCCCACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCCAC 1062
|||||||||||||||||||||||||||||||||||||||||||||||||||||
ATCCCACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCCAC 334

ACCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCAGTGATCTCGACTCC 1122
|||||||||||||||||||||||||||||||||||||||||||||||||||||
ACCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCAGTGATCTCGACTCC 274

GGCCACAGACCCCCAGGGCATTGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTC 1182
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGCCACAGACCCCCAGGGCATTGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTC 214

GACCCCACTTCAGGCACTAAGAGGGCTGGACCTGGCGGCGGCAGGAAGCCAAAGAGA 1242
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GACCCCACTTCAGGCACTAAGAGGGCTGGACCTGGCGGCGGCAGGAAGCCAAAGAGA 154

CCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAGCTTCCTCCCT 1302
||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   153   CTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTTCCTC
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy 1303   TGAGAATTCCCTGTGGATTTTTAAAACAGATATTATTTTTATTATTGTGACAAA
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   93   TGAGAATTCCCTGTGGATTTTTAAAACAGATATTTTTATTATTTTATTATTGTGACAAA

Qy 1363   TTGATAAATGG 1373
                |||||||||||
Db   33   TTGATAAATGG 23

RESULT 12
LOCUS   CF126539
DEFINITION UI-HF-ETO-avw-m-22-0-UI.r1 NIH_MGC_214 Homo sapiens cDNA cl
            IMAGE:30559317 5', mRNA sequence.
                                              666 bp   mRNA   linear   EST 05-
ACCESSION   CF126539
VERSION   CF126539.1   GI:33203878
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelec
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 666)
AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE   Normalization and subtraction: two approaches to facilitate
            discovery
JOURNAL   Genome Res. 6 (9), 791-806 (1996)
MEDLINE   97044477
PUBMED   8889548
COMMENT   Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156   MBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Mary Hendrix
            cDNA Library preparation: Dr. M. Bento Soares, University
            cDNA Library Arrayed by: Dr. M. Bento Soares, University c
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Distribution information can be found
            http://genome.uiowa.edu/distribution/humanfl.html
            Seq primer: pYX-5.
FEATURES   Location/Qualifiers
     source   1..666
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30559317"
                     /tissue_type="Chondrosarcoma Lung Metastasis cell
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_MGC_214"
                     /note="Organ: Lung; Vector: pYX-Asc; Site 1: EcoR
                     Site 2: Not I; The library was constructed accordi
                     Bonaldo, Lennon and Soares, Genome Research, 6:791
                     1996. Denatured RNA was size fractionated on a 1%
                     gel. First strand cDNA synthesis was primed with c
                     primer containing a Not I site. Double strand cDNA
                     size selected according to mRNA size fraction, lig
                     with EcoR I adaptor, digested with Not I and then
                     directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the po
                     is GATAAGCCA. Tissue was provided by Mary Hendrix

ORIGIN
Query Match   33.9%;   Score 465;   DB 14;   Length 666;
Best Local Similarity   99.6%;   Pred. No. 1.5e-225;
Matches 565;   Conservative   0;   Mismatches   2;   Indels   0;   Ga

Qy   303   GGACCCGTCGGAACTGAATCCCCAGACAGAAGAAAGCCAGGATCCTCGCGCCTTTCCT
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

High quality sequence stop: 451.

FEATURES   Location/Qualifiers
     source   1..474
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:6128001"
                  /tissue_type="Purified pancreatic islet"
                  /lab_host="DH10B"
                  /clone_lib="HR85 islet"
                  /note="Organ: Pancreas; Vector: pBluescript SK(-)
                  NotI; Site 2: XhoI; cDNA made by oligo-dT priming
                  Size-selected on agarose gel. Average insert size
                  XhoI site was destroyed after directional cloning
                  Amplified once. Contact information: Hiroshi Inou
                  Metabolism Div. (Alan Permutt Lab), Washington Un
                  School of Medicine, Box 8127, 660 South Euclid Av
                  Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu,
                  314-362-1916, Fax: 314-747-2692."

ORIGIN

Query Match   33.6%; Score 462; DB 13; Length 474;
Best Local Similarity 100.0%; Pred. No. 4.8e-224;
Matches 462; Conservative 0; Mismatches 0; Indels 0; C

Qy   912 TGGGCACCCGGTCCCCTCTGCCCACCCTCAGCCGCTCTTTGCTCCGAGACCTGCCC
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   474 TGGGCACCCGGTCCCCTCTGCCCACCCTCAGCCGCTCTTTGCTCCGAGACCTGCCC

Qy   972 CTCTAGAGGCTGCCTGGCCTGTTCAGTGTTTTCATCCCACATAAATACAGTA
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   414 CTCTAGAGGCTGCCTGGCCTGTTCAGTGTTTTCATCCCACATAAATACAGTA

Qy   1032 ACTCTTATCTTACAACTCCCCACCGCCCACTCTCCACCTCAGCTCCCCAAT
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   354 ACTCTTATCTTACAACTCCCCACCGCCCACTCTCCACCTCAGCTCCCCAAT

Qy   1092 ACCCTTTGAGGCCCCAGTAGTCTGACTCCCCCCTGGCCACAGAGACCCCAGGGC
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   294 ACCCTTTGAGGCCCCAGTAGTCTGACTCCCCCCTGGCCACAGAGACCCCAGGGC

Qy   1152 GTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCCCACTTCAGGCACTA
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   234 GTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCCCACTTCAGGCACTA

Qy   1212 GGCTGGACCTGGCGGCAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCA
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   174 GGCTGGACCTGGCGGCAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCA

Qy   1272 GAGGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAA
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   114 GAGGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAA

Qy   1332 ATATTATTTTTATTATTATTGTGACAAAATGTTGATAAATGG 1373
            ||||||||||||||||||||||||||||||||||||||||||
Db   54 ATATTATTTTTATTATTATTGTGACAAAATGTTGATAAATGG   13

RESULT 14
CB529199/c
LOCUS   CB529199   569 bp   mRNA   linear   EST 16
DEFINITION UI-H-FT2-bjf-k-03-0-UI.s1 NCI_CGAP_FT2 Homo sapiens cDNA c
            UI-H-FT2-bjf-k-03-0-UI 3', mRNA sequence.
ACCESSION   CB529199
VERSION   CB529199.1 GI:29389647
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 569)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE   National Cancer Institute, Cancer Genome Anatomy Project (

CCGTCGGAACTGAATCCCCAGACAGAAGAAGCCAGGATCCTCGCGCCTTCCTGAA 71

CTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAACACGGGCTCGAAGAGC 422
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAACACGGGCTCGAAGAGC 131

GCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACGGAGCGCAGGCAGGTGT 482
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACGGAGCGCAGGCAGGTGT 191

GGGACAGTGAGTGCTGGGAGGAAGCCAGAATCAAACAGCTCCAGCCCTTCGCGCTA 542
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGGACAGTGAGTGCTGGGAGGAAGCCAGAATCAAACAGCTCCAGCCCTTCGCGCTA 251

CGCCAGATCGGGAGTTTATAGTCACCCGGCTCTACACCTGTACTGTCA 602
||||||||||||||||||||||||||||||||||||||||||||||||
CGCCAGATCGGGAGTTTATAGTCACCCGGCTCTACACCTGTACTGTCA 311

CACTTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGGATGGTGT 662
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
CACTTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGGATGGTGT 371

GCCCTGCGCTGCCTGGAGGAATTCTCAGCCACTGCGCGCCAGTTCCCTCGGGCCCCA 722
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCCCTGCGCTGCCTGGAGGAATTCTCAGCCACTGCGCGCCAGTTCCCTCGGGCCCCA 431

CGCCTCTCGCCAGGTGTCTGGCGTTGGCCCTGCGCGGCCCAGGGTCCTCCCTGCGGAT 782
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
CGCCTCTCGCCAGGTGTCTGGCGTTGGCCCTGCGCGGCCCAGGGTCCTCCCTGCGGAT 491

ACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACTTCGGACTCTT 842
|||||||||||||||||||||||||||||||||||||||||||||||||||
ACCCTCCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACTTCGGACTCCTT 551

GTTCACTGAGGGGGCCCTGTCTC 869
|||||||||||||||||||||||
GTTCACTGAGGGGGCCCTGGTCTC 578

---

474 bp   mRNA   linear   EST 21-OCT-2002
...x1 HR85 islet Homo sapiens cDNA clone IMAGE:6128001 3',
sequence.
...5.1 GI:24203667

apiens (human)
sapiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 474)
,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
oka,I., Scearce,M., Brestell,J., Gradwohl,G., Clifton,S.,
r,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
t,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
as,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
ms,T., Jackson,Y. and Bowers,Y.
ine Pancreas Consortium
ished (2000)

t: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
ine Pancreas Consortium
d University, Howard Hughes Medical Institute
of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
38
17-495-1812
17-495-8557
dmelton@biohp.harvard.edu
y was constructed by Dr. Hiroshi Inoue DNA sequencing by:
gton University Genome Sequencing Center For information on
ing a clone please contact: Dr. Hiroshi Inoue
e@im.wustl.edu)
imer: -40UP from Gibco

ne Index
shed (1997)
: Robert Strausberg, Ph.D.
cgapbs-r@mail.nih.gov
Procurement: Dr. Gary W. Hunninghake, U of I
ibrary preparation: Dr. M. Bento Soares, University of Iowa
ibrary Arrayed by: Dr. M. Bento Soares, University of Iowa
quencing by: Dr. M. Bento Soares, University of Iowa
Distribution information can be found at
genome.uiowa.edu/distribution/cgap.html
lowing repetitive elements were found in this cDNA
e: 1-82, >AT rich#Low_complexity (matched compliment)
er: M13 FORWARD
es.

Location/Qualifiers
1..569
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FT2-bjf-k-03-0-UI"
/tissue_type="Aveolar Macrophage"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_FT2"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI_CGAP_FT2 is a subtracted cDNA library constructed from
a pool of 81 RNA samples from Alveolar Macrophages
challenged with different treatments. The library was
subtracted according to Bonaldo, Lennon and Soares, Genome
Research, 6:791-806, 1996. The tissue was provided by Dr.
Gary W. Hunninghake of the University of Iowa.
TAG_TISSUE=Human Lung Aveolar Macrophage
TAG_LIB=UI-H-FT2
TAG_SEQ=GGCCATGCCG"

rity   32.7%;   Score 449; DB 14; Length 569;
99.8%;   Pred. No. 2.1e-217;
nservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

GTCCCAGGCTGCCGGCTCCCCCTGCGACAGCTCTCTGGGCACCCGGTCCCCTGCC 933
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTCCCAGGCTGCCGGCTCCCCCTGCGACAGCTCTCTGGGCACCCGGTCCCCTGCC 468

CTCAGCGCGTCTTTGCTCCAGACCTGCCCCCTCCCTCTAGAGGCTGCCTGGGCCTG 993
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTCAGCGCGTCTTTGCTCCAGACCTGCCCCCTCCCTCTAGAGGCTGCCTGGGCCTG 408

GTGTTTTCCATCCCACCATAAATACACAGTATTCCCACTCTTATCTTACAACTCCCCC 1053
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTGTTTTCCATCCCACCATAAATACACAGTATTCCCACTCTTATCTTACAACTCCCCC 348

CCACTCTTCCACCTCACTAGCTCCCCAATCCCTGACCCCTTTGAGGCCCCCAGTGAT 1113
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCACTCTTCCACCTCACTAGCTCCCCAATCCCTGACCCCTTTGAGGCCCCCAGTGAT 288

CTCCCCCCTGGCCACAGACCCCCACCCCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCAA 1173
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTCCCCCCTGGCCACAGACCCCCACCCCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCAA 228

GGTCCAGAAGACCCCCACTTCAGGCACTTAAGGAGGGGGCTGGACCTGGGCGCGCAGGAAG 1233
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGTCCACAAGACCCCCACTTCAGGCACTTAAGGAGGGGGCTGGACCTGGGCGCGCAGGAAG 168

GAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAG 1293
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAG 108

CCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTATTTTTATTATTATTGT 1353
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTATTTTTATTATTATTGT 48

Qy   1354 GACAAAATGTTGATAAATGG 1373
             ||||||||||||||||||||
Db   47 GACAAAATGTTGATAAATGG 28

RESULT 15
BQ707185
LOCUS   BQ707185
DEFINITION   AGENCOURT_8353983 NIH_MGC_113 Homo sapiens cDNA clone IMAGE
             5', mRNA sequence.
ACCESSION   BQ707185
VERSION   BQ707185.1 GI:21846084
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleo
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 948)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (M
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: Dr. Mark Watson
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLCM2466 row: n column: 17
             High quality sequence start: 24
             High quality sequence stop: 550.
FEATURES   Location/Qualifiers
     source   1..948
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6278608"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_113"
                     /note="Organ: spleen; Vector: pOTB7; Site 1: XhoI;
                     EcoRI; cDNA made by oligo-dT priming. Directionall
                     into EcoRI/XhoI sites using the following 5' adapt
                     GGCACGAG(G). Library constructed by Ling Hong in
                     laboratory of Gerald M. Rubin (University of Cali
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene
                     Superscript II RT (Life Technologies). Note: this
                     NIH_MGC Library."

ORIGIN

Query Match   31.8%;   Score 437; DB 13; Length 948;
Best Local Similarity 99.8%; Pred. No. 3.2e-211;
Matches 487; Conservative 0;   Mismatches 1;   Indels 0;   Ga

Qy   475 GCAGGTGTGGACGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAGC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   167 GCAGGTGTGGACGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAGC

Qy   535 CTTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTGCACCCGGGCTGGGCTCTACTAC
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   227 CTTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTGCACCCGGGCTGGGCTCTACTAC

Qy   595 TACTGTCAGGTGCACTTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   287 TACTGTCAGGTGCACTTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTG

Qy   655 GATGGTGTGCTGCCCTGCGCTGCGTGGGGAGGAATTCTCAGCCACTGCGGCCAGTCC
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   347 GATGGTGTGCTGCCCTGCGCTGCGTGGGGAGGAATTCTCAGCCACTGCGGCCAGTTCC

Qy   715 GGGCCCCAGCTCCGCCTCTGCCAGGTGTCTGGGCTGTGGCCCTGCGGGCCCAGGGTCC
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Qy  301  CAGGACCCGTCGGAACTGAATCCCCAGACAGAGAAAGCCAGGATCCTCGCGCCTTT
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  CAGGACCCGTCGGAACTGAATCCCCAGACAGAGAAAGCCAGGATCCTCGCGCCTTT

Qy  361  AACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAACACGGGCTCG
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  AACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAACACGGGCTCG

Qy  421  GCGATGCGAGCCCATTATGAAGTTCATCCACGACCTTGACACGAGGCGGAGGCGGAGGC
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  GCGATGCGAGCCCATTATGAAGTTCATCCACGACCTTGACACGAGGCGGAGGCGGAGGC

Qy  481  GTGGACGGACACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTTCAGCCCTCT
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  GTGGACGGACACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTTCAGCCCTCT

Qy  541  TACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTTACTACCTGTA
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  TACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTTACTACCTGTA

Qy  601  CAGGTGCACTTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGGA
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  CAGGTGCACTTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGGA

Qy  661  GTGCTGGCCCTGCGCTGCCTGGAGAATTCTCAGCCACCTGCGGCCAGTTCCCTCGG
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  GTGCTGGCCCTGCGCTGCCTGGAGAATTCTCAGCCACCTGCGGCCAGTTCCCTCGG

Qy  721  CAGCTCCGCCTCTGCC  736
              ||||||||||||||||
Db  421  CAGCTCCGCCTCTGCC  436
```

RESULT 17

CF126932

LOCUS   CF126932   697 bp   mRNA   linear   EST 05
DEFINITION   UI-HF-ET0-avx-k-19-0-UI.r1 NIH_MGC_214 Homo sapiens cDNA c
             IMAGE:30563490 5', mRNA sequence.
ACCESSION   CF126932
VERSION   CF126932.1   GI:33204664
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 697)
AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE   Normalization and subtraction: two approaches to facilitat
             discovery
JOURNAL   Genome Res. 6 (9), 791-806 (1996)
MEDLINE   97044477
PUBMED   8889548
COMMENT   Contact: Soares, MB
             Coordinated Laboratory for Computational Genomics
             University of Iowa
             375 Newton Road, 4156   MEBRF, Iowa City, IA 52242, USA
             Tel: 319 335 8250
             Fax: 319 335 9565
             Email: bento-soares@uiowa.edu
             Tissue Procurement: Mary Hendrix
             cDNA Library preparation: Dr. M. Bento Soares, University
             cDNA Library Arrayed by: Dr. M. Bento Soares, University
             DNA Sequencing by: Dr. M. Bento Soares, University of Iow
             Clone Distribution: Distribution information can be found
             http://genome.uiowa.edu/distribution/humanfl.html
             The following repetitive elements were found in this cDN
             sequence: 37-143, >GC_rich#Low_complexity
             Seq primer: pYX-5.
FEATURES   Location/Qualifiers
     source   1..697
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"

```
Qy   ??  CCCAGCTCCGCCTCTGCCAGGTGTCTGCGGCTCGTGTTGGCCTGCGGCCAGGGTCCTCC  466
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  834

Qy  526  GGATCCGCACCCCCCGGGCCATCTCAAGGCTGCCCCCTTCCTCACCTACTTC
              ||||||||||||||||||||||||||||||||||||||||
Db  526  GGATCCGCACCCCCCGGGCCATCTCAAGGCTGCCCCCTTCCTCACCTACTTC

Qy  894  TCTTCCAGGTTCACTGAGGGGCCCTGGTCTCCCACAGTCGTCCCAGGCTGCCGGC
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      TCTTCCAGGTTCACTGAGGGGCCCTGGTCTCCCACAGTCGTCCCAGGCTGCCGGC  586

Qy  954  CTCGACAGCTCTCTGGGCACCCGGTCCCCGGTCCCCTTGCCCCACCCTCAGCCGCTCTTTGC
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  646  CTCGACAGCTCTCTGGGCACCCGGTCCCCGGTCCCCTTGCCCCACCCTCAGCCGCTCTTTGC

Qy  962  GACC
         ||||
Db  654  GACC
```

31
URT_8682031 Lupski_sciatic_nerve Homo sapiens cDNA clone   EST 16-AUG-2002
6197488 5', mRNA sequence.
31
31.1   GI:22276239

sapiens (human)
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 940)

C http://mgc.nci.nih.gov/.
al Institutes of Health, Mammalian Gene Collection (MGC)
ished (1999)
t: Robert Strausberg, Ph.D.
cgapbs-r@mail.nih.gov
Procurement: Dr. James R. Lupski
Library Preparation: Life Technologies, Inc.
Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
equencing by: Agencourt Bioscience Corporation
distribution: MGC clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL at:
/image.llnl.gov
LLAM13607 row: j column: 17
uality sequence stop: 453.
         1..940
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6197488"
                     /clone_lib="Lupski_sciatic_nerve"
                     /sex="male"
                     /tissue_type="sciatic nerve"
                     /dev_stage="adult, 70 yr"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCACGGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.87 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life
                     Technologies."

arity 100.0%; Pred. No. 1e-210;
onservative 0; Mismatches 0; Indels 0; Gaps 0;
31.8%; Score 436; Pred. No. 1e-210; DB 13; Length 940;

/clone="IMAGE:30563490"
/tissue_type="Chondrosarcoma Lung Metastasis cell lines"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_MGC_214"
/note="Organ: Lung; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with Not I and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is GATAAGGCCA. Tissue was provided by Mary Hendrix."

FEATURES
source   1..834
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:5206217"
              /lab_host="DH10B"
              /clone_lib="NIH_MGC_122"
              /note="Organ: pooled lung and spleen; Vector: pCM
              Site 1: NotI; Site 2: EcoRV (destroyed); RNA sourc
              anonymous pool of 24 week female lung, 16 week fem
              spleen, and 20-22 week male spleens. Library is c
              primed and directionally cloned (EcoRV site is des
              upon cloning). Average insert size 1.4 kb, inser
              range 1-3 kb. Library is normalized and enriched f
              full-length clones and was constructed by C. Grub
              (Invitrogen).   Research Genetics tracking code 02
              this is a NIH_MGC Library."

ORIGIN

Query Match   30.3%;   Score 416;   DB 12;   Length 834;
Best Local Similarity 99.7%;   Pred. No. 1.6e-200;
Matches 656; Conservative   0; Mismatches   0;   Indels   2;   Ga

Qy   272 CCCAGGAGGAGCTGGTGGCAGAGGAGGACCAGGACCCGTCGGAACTGAATCCCCAGA
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CCCAGGAGGAGCTGGTGGCAGAGGAGGACCAGGACCCGTCGGAACTGAATCCCCAGA

Qy   332 AAGAAAGCCAGATCCTGCGCCTTCTGAACCGACTAGTTCGGCCTCGCAGAAGTC
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AAGAAAGCCAGATCCTGCGCCTTCTGAACCGACTAGTTCGGCCTCGCAGAAGTC

Qy   392 CTAAAGGCCGGAAAAACACGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATC
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CTAAAGGCCGGAAAAACACGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATC

Qy   452 GACC-TGGACAGGACGGAGCGCAGGCAGGCGCAGGCGCAGGTGTGACGGGACAGTGGCTGGGAC
           |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GACCGTGGACAGGACGGAGCGCAGGCAGGCGCAGGTGTGACGGGACAGTGGCTGGGAC

Qy   511 GCCAGAATCAACAGCTCCAGCCCTCGCTACAACCGCCAGATCGGGAGTTTATA
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GCCAGAATCAACAGCTCCAGCCCTCGCTACAACCGCCAGATCGGGAGTTTATA

Qy   571 ACCCGGCTGGGCTCTACACTTGCTACTTGTCAGTGCACTTGATGAGGGGAAGGCT
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ACCCGGCTGGGCTCTACACTTGCTACTTGTCAGTGCACTTTGATGAGGGGAAGGCT

Qy   631 TACCTGAAGCTGACTTGCTGGTGGATGGTGTGCTGGCCCTGGTGTCTGGG
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TACCTGAAGCTGACTTGCTGGTGGATGGTGTGCTGGCCCTGGTGTCTGGG

Qy   691 TCAGCCACTGCGCGCCAGTTCCCTCGGCGCCCAGCTCCAGCTCCGCCTCTGCCAGTGTCTGGG
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TCAGCCACTGCGCGCCAGTTCCCTCGGCGCCCAGCTCCAGCTCCGCCTCTGCCAGTGTCTGGG

Qy   751 TTGGCCCTGCGGCCAGGGTCCTCCTGCGGATCCGCACCCTCCCTGGGCCCATCTC
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TTGGCCCTGCGGCCA-GGTCCTCCTGCGGATCCGCACCCTCCCTGGGCCCATCTC

Qy   811 GCTGCCCCTTCCTCACCTACTTCGGACTTCCAGGTTCACTGAGGGCCCTGGTC
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   540 GCTGCCCCTTCCTCACCTACTTCGGACTTCCAGGTTCACTGAGGGCCCTGGTC

Qy   871 CCCAGAGTCGTCCCAGGCTGCCAGCTCCCCTGCCACAGCTTCTGTGGGCACCCGGTCCCC

31.0%;   Score 425;   DB 14;   Length 697;
rity 100.0%;   Pred. No. 4e-205;
nservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

GCCTGCCCAGGAGGAGCTGGTGGCAGAGGAGGACCCGTCGGAACTGAAT   321
|||||||||||||||||||||||||||||||||||||||||||||||||
GCCTGCCCAGGAGGAGCTGGTGGCAGAGGAGGACCCGTCGGAACTGAAT   268

GACAGAGAAAGCCAGATCCTGCGCCTTCTGAACCGACTAGTTCGGCCTCGC   381
|||||||||||||||||||||||||||||||||||||||||||||||||||
GACAGAGAAAGCCAGATCCTGCGCCTTCTGAACCGACTAGTTCGGCCTCGC   328

TGCACCTAAAGGCCGGAAAAACACGGGCTCGAAGAGCGATCGCAGCCCATTATGAA   441
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGCACCTAAAGGCCGGAAAAACACGGGCTCGAAGAGCGATCGCAGCCCATTATGAA   388

TCCACGACCTGACAGGACGGAGCGCAGGCAGCGGGACAGTGTGGACGGGACAGTGAGTGGC   501
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCCACGACCTGACAGGACGGAGCGCAGGCAGCGGGACAGTGTGGACGGGACAGTGAGTGGC   448

GGAAGCCAGAATCAACAGCTCCAGCCCTCGGCCCTCCAGCCGCCACAACCGCCAGATCGGGGAG   561
||||||||||||||||||||||||||||||||||||||||||||||||||
GGAAGCCAGAATCAACAGCTCCAGCCCTCGGCCCTCCAGCCGCCACAACCGCCAGATCGGGGAG   508

AGTCACCCGGCTGGGCTCTACTACCTTGTCACTGTCAGTGCACTTTGATGAGGGG   621
||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGTCACCCGGCTGGGCTCTACTACCTTGTCACTGTCAGTGCACTTTGATGAGGGG   568

TGTCTACCTGAAGCTGGACTTGCTGGTGGATGGTGTGCTGGCCCTGCGCTGCCTG   681
||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGTCTACCTGAAGCTGGACTTGCTGGTGGATGGTGTGCTGGCCCTGCGCTGCCTG   628

686

633

6   834 bp   mRNA   linear   EST 25-SEP-2001
66F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5206217 5',
quence.
6
6.1   GI:15758344

piens (human)
piens
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 834)
http://mgc.nci.nih.gov/.
l Institutes of Health, Mammalian Gene Collection (MGC)
shed (1999)
: Robert Strausberg, Ph.D.
cgapbs-r@mail.nih.gov
Procurement: Life Technologies, Inc.

---

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information car
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11517 row: c   column: 18
High quality sequence stop:772.

Db   233 TACTCTGTGGCAAGGATGGTGTCCAGAAGAGACCCCACTCACTTCAGGCACTCAAGAGAGGGGCT
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy 1220 CTGGCCGCAGGAAGCCAAAGAGACTGGGCTCAGGCAGGAGTTCCCAAATGTGAGG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   173 CTGGCCGCAGGAAGCCAAAGAGACTGGGCTCAGGCAGGAGTTCCCAAATGTGAGG

Qy 1280 AGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTAAAACAGATAT
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   113 AGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTAAAACAGATAT

Qy 1340 TTTATTATTATTGTGACAAAATGTTGATAAATGG 1373
          ||||||||||||||||||||||||||||||||||
Db   53 TTTATTATTATTGTGACAAAATGTTGATAAATGG 20

RESULT 20
BM971606/c
LOCUS   BM971606   568 bp   mRNA   linear   EST 20
DEFINITION UI-CF-EC1-abl-p-06-0-UI.s1 UI-CF-EC1 Homo sapiens cDNA clo
            UI-CF-EC1-abl-p-06-0-UI 3', mRNA sequence.
ACCESSION   BM971606
VERSION   BM971606.1 GI:19589193
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 568)
AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE   Normalization and subtraction: two approaches to facilitat
            discovery
JOURNAL   Genome Res. 6 (9), 791-806 (1996)
MEDLINE   97044477
PUBMED   8889548
COMMENT   Contact: McCray, PB
            McCray Lab
            University of Iowa
            2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
            Tel: 319 356 4866
            Fax: 319 356 7171
            Email: paul-mccray@uiowa.edu
            Tissue Procurement: Dr. M. J. Welsh, University of Iowa
            cDNA Library preparation: Dr. M. Bento Soares, University
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iow
            DNA Sequencing by: Dr. M. Bento Soares, University of Iow
            Clone Distribution: Researchers may obtain clones from Re
            Genetics (www.resgen.com) or from Open Biosystems
            (www.openbiosystems.com).
            The following repetitive elements were found in this cDN/
            sequence: 1-82, >AT_rich#Low_complexity (matched complimer
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES   Location/Qualifiers
     source   1..568
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-CF-EC1-abl-p-06-0-UI"
                     /tissue_type="Lung"
                     /dev_stage="Adult and Fetal"
                     /lab_host="DH10B (Life Technologies) (T1 phage re
                     /clone_lib="UI-CF-EC1"
                     /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia)
                     modified polylinker; Site 1: EcoR I; Site 2: Not
                     UI-CF-EC1 is a normalized cDNA library containing
                     following tissue(s): Normal lung from adult and
                     day 64, day 87, week 19 and week 42. The library
                     constructed according to Bonaldo, Lennon and Soar
                     Genome Research, 6:791-806, 1996. First strand cr
                     synthesis was primed with an oligo-dT primer cont
                     Not I site. Double stranded cDNA was ligated to
                     adaptor, digested with Not I, and cloned directio

---

Db   GTCGTCCAGGCTGCCGGCTCTCCCCTTGACAGCTCTCTCGGGCACCCGGTCCCCT 657
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||

96   6.x1 NCI_CGAP_Brn23   413 bp   mRNA   linear   EST 30-MAR-1999
      6.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:2096962 3',
      sequence.
96   GI:4268727
96.1 GI:4268727

            apiens (human)
            apiens
            ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            ses 1 to 413)
            NDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            al Cancer Institute / National Institute of Neurological
            ers and Stroke, Brain Tumor Genome Anatomy Project
            BTGAP), Tumor Gene Index
            ished (1998)
            t: Robert Strausberg, Ph.D.
            cgapbs-r@mail.nih.gov
            Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,

            Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            o, Ph.D.
            Library Arrayed by: Greg Lennon, Ph.D
            equencing by: Washington University Genome Sequencing Center
            distribution: NCI-CGAP clone distribution information can be
            through the I.M.A.G.E. Consortium/LLNL at:
            o.llnl.gov/bbrp/image/image.html
            imer: -40UP from Gibco
            Length: 1728   Std Error: 0.00
            quality sequence stop: 410.
            Location/Qualifiers
            1..413
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2096962"
                     /tissue_type="glioblastoma (pooled)"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Brn23"
                     /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCGAAGTGGAGCGGCCGCATATCTTTTTTTTTTTTTTTTTTTTTTTT
                     T 3'); double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library is normalized, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."

arity   28.7%; Score 394; DB 9; Length 413;
            100.0%; Pred. No. 2.4e-189;
conservative   0; Mismatches   0; Indels   0; Gaps   0;

CCTGGCCGTGTTCACGTGTTTCATCCCACATAAATACAGTATTCCACTCTTAT 1039
||||||||||||||||||||||||||||||||||||||||||||||||||||
CCTGGCCGTGTTCAGTGTTTCATCCCACATAAATACAGTATTCCCACTCTTAT 354

ACAACTCCCCCACCGCCCACTTCCACCTCCAGCTTCCCAATCCCTGACCCTTTG 1099
||||||||||||||||||||||||||||||||||||||||||||||||||||
ACAACTCCCCCACCGCCCACTCTCACCTCACTAGCTCCCAATCCCTGACCCTTTG 294

CCCCAGTGATCTCGACTCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTG 1159
||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCCAGTGATCTCGACTCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTG 234

CCTGTGTGGGCAAGGATGGTCCAGAAGAGACCCCACTTCAGGCACTAAGAGGGGCTCGAC 1219
||||||||||||||||||||||||||||||||||||||||||||||||||

FEATURES   Location/Qualifiers
source   1..569
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="UI-H-FE1-bdu-c-24-0-UI"
                  /tissue_type="cell lines"
                  /dev_stage="Adult"
                  /lab_host="DH10B (Life Technologies)"
                  /clone_lib="NCI_CGAP_FE1"
                  /note="Organ: Chondrosarcoma; Vector: pT7T3-Pac
                  (Pharmacia) with a modified polylinker; Site 1: Ec
                  Site 2: Not I; NCI_CGAP_FE1 is a normalized cDNA
                  derived from a pool of mRNA obtained from 3 cell l
                  from grade II chondrosarcoma tissues. The library
                  constructed according to Bonaldo, Lennon and Soare
                  Genome Research, 6:791-806, 1996. First strand cDN
                  synthesis was primed with an oligo-dT primer conta
                  Not I site. Double stranded cDNA was ligated to a
                  adaptor, digested with Not I, and cloned directio
                  into pT7T3-Pac vector. The oligonucleotide used to
                  the synthesis of first-strand cDNA contains a libr
                  sequence that is located between the Not I site an
                  (dT)18 tail. The sequence tag for this library is
                  CGCTACGGAC. The cell lines were provided by Dr Jan
                  Martin from the University of Iowa.
                  TAG_TISSUE=Human grade 2 chondrosarcoma cell line
                  TAG_LIB=UI-H-FE1
                  TAG_SEQ=CGCTACGGAC"

ORIGIN

     Query Match   28.6%; Score 393; DB 13; Length 569;
     Best Local Similarity 99.8%;   Pred. No. 8.2e-189;
     Matches 443; Conservative 0; Mismatches 1; Indels 0; Ga

Qy   930 TGCCCCACCCTCAGCGCGCTCTTGCTCCAGACCTGCCCCTCCCCTCTCAGAGGCTGCC
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   471 TGCCCCACCCTCAGCGCGCTCTTGCTCCAGACCTGCCCCTCCCCTCTCAGAGGCTGCC

Qy   990 CCTGTTCAGTGTTTCATCCCACATAAATACAGTATTCCCACTCTTATCTTACA
               |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   411 CCTGTTCAGTGTTTCATCCCACATAAATACAGTATTCCCACTCTTATCTTACA

Qy   1050 CCCACCGCCCACTCTCCACCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCC(
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   351 CCCACCGCCCACTCTCCACCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCC(

Qy   1110 TGATCTGACTCCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCT(
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   291 TGATCTGACTCCCCCCCTGGCCACAGACCCCCAGGGTCATTGTGTTCACTGTACTCT(

Qy   1170 GCAAGGATGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGG
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   231 GCAAGGATGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGG

Qy   1230 GAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACJ
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   171 GAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACJ

Qy   1290 CAAGGTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTATTTTTATTJ
               |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   111 CAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTATTTTTATTJ

Qy   1350 TTGTGACAAAATGTTGATAAATGG 1373
               ||||||||||||||||||||||||
Db   51 TTGTGACAAAATGTTGATAAATGG 28

RESULT 22
BI824443
LOCUS   BI824443   531 bp   mRNA   linear   EST 04-
DEFINITION   603038693F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5179!
             mRNA sequence.

Qy   1187 AGACCC 1192
           ||||||
Db   526 AGACCC 531

RESULT 23
LOCUS   CB998034
DEFINITION   AGENCOURT_13903807 NIH_MGC_148 Homo sapiens cDNA clone
             IMAGE:30348032 5', mRNA sequence.
ACCESSION   CB998034
VERSION   CB998034.1 GI:30292554
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 824)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (
             MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: Dr. Stefan Hansson
             cDNA Library Preparation: Michael J. Brownstein (NHGRI) w
             and advice from Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information ca
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: NDAM390 row: b column: 09
             High quality sequence stop: 415.
FEATURES   Location/Qualifiers
     source   1..824
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30348032"
                     /tissue_type="pre-eclamptic placenta"
                     /lab_host="DH10B TonA"
                     /clone_lib="NIH_MGC_148"
                     /note="Organ: placenta; Vector: pBluescriptR; Sit
                     alI-XhoI; Site 2: BamH; Library is oligo-dT prime
                     directionally cloned using primer
                     5'-TTTTTTTTTTTTTTVN-3', size selected for avera
                     size 2.3 kb and normalized to ROT 5. This is a p
                     library enriched for full-lenght clones and cons
                     using the Cap-trapper method (Carninci, in prepa
                     Library constructed by M. Brownstein (NIMH/NHGRI
                     National Institutes of Health). Note: this is a N
                     Library."

ORIGIN
     Query Match   26.4%;   Score 362; DB 14; Length 824;
     Best Local Similarity   99.8%;   Pred. No. 5.9e-173;
     Matches   412;   Conservative   0;   Mismatches   1;   Indels   0;   C

Qy   52 TCCCTCGGGTCCCGGATGGGGGGCGTGAGCGAGGCAGGCACAGCGCCCCCGCCCCCAT
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   70 TCCCTCGGGTCCCGGATGGGGGGCGTGAGCGAGGCAGGCACAGCGCCCCCGCCCCCAT

Qy   112 GCCCGTCGGAGCCAGAGGCGGAGGGGGGCGCCGGGGGGAGCCGGCCACCGCCCTGCT
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   130 GCCCGTCGGAGCCAGAGGCGGAGGGGGGCGCCGGGGGGAGCCGGCCACCGCCCTGCT

Qy   172 CCGGCTCCGCGTCGGGCTCGGGCCTGGGCCTGCTGGGCCTCCTGCTGGCCGT
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   190 CCGCTCCGCGTCGGGCTCGGGCCTGGGCCTGCTGGGCCTCCTGCTGGCCGT

Qy   232 AGTTTGGGGAGCCGGGCATGCGTGTCCCCCAGGCAAGGATGGGTCCAGA
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   525 ...

13
13.1 GI:15935993

apiens (human)
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 531)
: http://mgc.nci.nih.gov/.
al Institutes of Health, Mammalian Gene Collection (MGC)
ished (1999)
t: Robert Strausberg, Ph.D.
cgapbs-r@mail.nih.gov
Procurement: Life Technologies, Inc.
Library Preparation: Life Technologies, Inc.
Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
equencing by: Incyte Genomics, Inc.
distribution: MGC clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL at:
/image.llnl.gov
LLAM11447 row: j column: 23
uality sequence stop: 529.
                 Location/Qualifiers
                 1..531
                 /organism="Homo sapiens"
                 /mol_type="mRNA"
                 /db_xref="taxon:9606"
                 /clone="IMAGE:5179510"
                 /lab_host="DH10B"
                 /clone_lib="NIH_MGC 115"
                 /note="Organ: pooled brain, lung, testis; Vector:
                 pCMV-SPORT6; Site 1: NotI; Site_2: EcoRV (destroyed); RNA
                 source anonymous Pool of 6 male brains, age range 23-27; 1
                 male lung, age 27; and 1 male testis, age 69. Library is
                 oligo-dT primed and directionally cloned (EcoRV site is
                 destroyed upon cloning). Average insert size 1.8 kb,
                 insert size range 1-3 kb library is normalized and
                 enriched for full-length clones and was constructed by C.
                 Gruber (Invitrogen) Research Genetics tracking code
                 021. Note: this is a NIH_MGC Library."

     arity   27.3%;   Score 375; DB 12; Length 531;
     onservative   0;   Pred. No. 1.2e-179;
                       0; Mismatches   1; Indels   1; Gaps   0;

CTCCCTCGCGGATCCGACACCTCCCTCGGGCCATCTCAAGGCTGCCCCCCTTCCTCA 826
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTCCCTCGCGGATCCGACACCTCCCTCGGGCCATCTCAAGGCTGCCCCCCTTCCTCA 165

CTTCGGACTCTTCCAGGTTCACTGAGGGCCCTGGTCTCCCCACAGTCGTCCCAGG 886
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTTCGGACTCTTCCAGGTTCACTGAGGGCCCTGGTCTCCCCGCAGTCGTCCCAGG 225

CGGCTCCCCTGACAGCTCTCTGGGCACACCGGTCCCCGGTCCCCTGCCCCACCCTCAGCCG 946
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CGGCTCCCCTGACAGCTCTCTGGGCACACCGGTCCCCGGTCCCCTGCCCCACCCTCAGCCG 285

TTGCTCCAGACCTGCCCCTCCCTCTAGAGGCTGCCTGGGCCTGTTCACGTGTTTTC 1006
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TTGCTCCAGACCTGCCCCTCCCTCTAGAGGCTGCCTGGGCCTGTTCACGTGTTTTC 345

CCACATAAATACAGTATTCCCACTCTTATTCTTACAACTCCCCCACCGCCCACTCTC 1066
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCACATAAATACAGTATTCCCACTCTTATTCTTACAACTCCCCCACCGCCCACTCTC 405

TCACTAGCTCCCCAATCCCTGACCCCTTTGAGGCCCCCAGTGATCTCGACTCCCCCC 1126
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCACTAGCTCCCCAATCCCTGACCCCTTTGAGGCCCCCAGTGATCTCGACTCCCCCC 465

CACAGACCCCCAGGGCATTGTGTTCACTGTACTTCGTGGGCAAGGATGGGTCCAGA 1186
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CACAGACCCCCAGGGCATTGTGTTCACTGTACTTCGTGGGCAAGGATGGGTCCAGA 525

20. Single-stranded (unhybridized) plasmids were by hydroxyapatite chromatography and used to make library."

ORIGIN

```
Query Match       25.6%;   Score 352;  DB 12;   Length 440;
Best Local Similarity 99.8%;  Pred. No. 6.6e-168;
Matches 402;  Conservative 0;  Mismatches 1;  Indels  0;  Ga

Qy  808 AAGGCTGCCCCTTCCTCACTTACTTCGGACTCTTCCAGGTTCACTGAGGGGCCCTG
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   38 AAGGCTGCCCCTTCCTCACTTACTTCGGACTCTTCCAGGTTCACTGAGGGGGCCTG

Qy  868 TCCCCACAGTCGTCCCAGGCTGCCGGCTCCCGGCTCCCGGCACCGGTC
            |||||||||||||||||||||||||||||||||||||||||||||||
Db   98 TCCCCACAGTCGTCCCAGGCTGCCGGCTCCCGGCTCCCGGCACCCGGTC

Qy  928 TCTGCCCCACCCCAGCCGCTCTTGCTCCAGACCTGCCCCTCCCTCAGAGGCTGC
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  158 TCTGCCCCACCCTCAGCCGCTCTTGCTCCAGACCTGCCCCTCCCTCAGAGGCTGC

Qy  988 GGCCTGTTCAGTGTTTTCCATCCCACATAAATACAGTATTCCCACTCTTATCTTAC
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  218 GGCCTGTTCAGTGTTTTCCATCCCACATAAATACAGTATTCCCACTCTTATCTTAC

Qy 1048 TCCCCACCGCCCACTCTTCCACCTCACTAGCTCCCCAATCCTGACCCTTTGAGGCC
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  278 TCCCCACCGCCCACTCTTCCACCTCACTAGCTCCCCAATCCTGACCCTTTGAGGGC

Qy 1108 AGTGATCTCGACTCCCCCCTGGCCACAGACCCCCAGGGCATTGTTCACTGTACTC
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  338 AGTGATCTCGACTCCCCCCTGGCCACAGACCCCCAGGGCATTGTTCACTGTACTC

Qy 1168 GGGCAAGGATGGGTCCAGAAGACCCCACTTCAGGCACTAAGAG 1210
            |||||||||||||||||||||||||||||||||||||||||||
Db  398 GGGCAAGGATGGGTCCAGAAGACCCCACTTCAGGCACTAAGAG 440
```

RESULT 25

BQ674188

LOCUS   BQ674188   951 bp   mRNA   linear   EST 15-

DEFINITION AGENCOURT_8354100 NIH_MGC_102 Homo sapiens cDNA clone IMAGI
            5', mRNA sequence.

ACCESSION   BQ674188

VERSION   BQ674188.1 GI:21785022

KEYWORDS   EST.

SOURCE   Homo sapiens (human)

ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleo
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE   1 (bases 1 to 951)

AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.

TITLE   National Institutes of Health, Mammalian Gene Collection (b

JOURNAL   Unpublished (1999)

COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information car
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2459 row: d column: 01
            High quality sequence stop: 335.

FEATURES   Location/Qualifiers

source   1..951
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:6275664"
                /tissue_type="epidermoid carcinoma, cell line"
                /lab_host="DH10B (phage-resistant)"

FEATURES   Seq primer: -40RP from Gibco.
source   Location/Qualifiers
                  1..377
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /tissue_type="insulinoma"
                  /lab_host="DH10B (phage-resistant)"
                  /clone_lib="Human insulinoma"
                  /note="Organ: pancreas; Vector: pBluescript SK-;
                  XhoI; Site_2: EcoRI; Constructed with lambda ZAPI
                  (Stratagene) by Dr. J. Ferrer, in vivo mass-excis
                  pBluescript SK- by Dr. H. Inoue following the Wa
                  University protocol
                  (http://genome.wustl.edu/est/lambda_protocol.shtm
                  Please contact Hiroshi Inoue, MD/PhD for further
                  information on this library (Metabolism Division
                  Laboratory, Washington University School of Medi
                  8127, 660 S Euclid Ave, St. Louis, MO 63110). No
                  is a Washington University Pancreas EST project 1

ORIGIN

                  Query Match   24.8%;   Score 341;   DB 12;   Length 377;
                  Best Local Similarity   100.0%;   Pred. No. 2.6e-162;
                  Matches 341;   Conservative   0;   Mismatches   0;   Indels   0;   C

Qy   1033 CTCTTATCTTACAACTCCCCACCGCCCACTCTCCACCTCACTAGCTCCCAATCC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8 CTCTTATCTTACAACTCCCCACCGCCCACTCTCCACCTCACTAGCTCCCAATCC

Qy   1093 CCCTTTGAGGCCCCCAGTGATCTCGACTCCCCCCTGGCCACAGACCCCCAGGGCAT
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   68 CCCTTTGAGGCCCCCAGTGATCTCGACTCCCCCCTGGCCACAGACCCCCAGGGCAT

Qy   1153 TTCACTGTACTCTGTGGGCAAGGATGGTGGTCCAAGAGACCCCACTTCAGGCACTAAC
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   128 TTCACTGTACTCTGTGGGCAAGGATGGTGGTCCAAGAGACCCCACTTCAGGCACTAAC

Qy   1213 GCTGGACCTGGCGCGGCAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAA
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188 GCTGGACCTGGCGCGGCAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAA

Qy   1273 AGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTAAA
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   248 AGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTAAA

Qy   1333 TATTATTTTTATTATTATTGTGACAAAATGTTGATAAATGG 1373
           ||||||||||||||||||||||||||||||||||||||||||
Db   308 TATTATTTTTATTATTATTGTGACAAAATGTTGATAAATGG 348

RESULT 27
BE858778/c
LOCUS   BE858778   710 bp   mRNA   linear   EST 2!
DEFINITION 7f95b06.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:3!
            similar to contains element MER32 repetitive element .; m!
            sequence.
ACCESSION   BE858778
VERSION   BE858778.1 GI:10374165
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteli
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 710)
  AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE   National Cancer Institute / National Institute of Neuroloc
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGAP), Tumor Gene Index
  JOURNAL   Unpublished (1998)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfe

TITLE   Normalization and subtraction: two approaches to facilitate
           discovery
JOURNAL   Genome Res. 6 (9), 791-806 (1996)
MEDLINE   97044477
PUBMED   8889548
COMMENT   Contact: Soares, MB
           Coordinated Laboratory for Computational Genomics
           University of Iowa
           375 Newton Road   4156   MEBRF, Iowa City, IA 52242, USA
           Tel: 319 335 8250
           Fax: 319 335 9565
           Email: bento-soares@uiowa.edu
           Tissue Procurement: Dr. Gregg Hageman
           cDNA Library preparation: Dr. M. Bento Soares, Univeristy
           cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy c
           DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
           Clone Distribution: Researchers may obtain clones from Res
           Genetics (www.resgen.com).
           The following repetitive elements were found in this cDNA
           sequence: 1-94, >AT rich#Low_complexity (matched compliment
           Seq primer: M13 Forward
           POLYA=Yes.
FEATURES   Location/Qualifiers
     source   1..367
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-CQ1-acs-a-05-0-UI"
                     /tissue type="optic nerve"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage res
                     /clone_lib="UI-E-CQ1"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) v
                     modified polylinker; Site_1: EcoR I; Site_2: Not I
                     UI-E-CQ1 is a normalized cDNA library containing t
                     following tissue(s): optic nerve. The library was
                     constructed according to Bonaldo, Lennon and Soars
                     Genome Research, 6:791-806, 1996. First strand cDI
                     synthesis was primed with an oligo-dT primed conta
                     Not I site. Double stranded cDNA was ligated to ai
                     adaptor, digested with Not I, and cloned direction
                     into pT7T3-Pac vector. The oligonucleotide used tc
                     the synthesis of first-strand cDNA contains a libi
                     sequence that is located between the Not I site ai
                     (dT)18 tail. The sequence tag for this library is
                     CCATTAAGTG. This library was created for the prog
                     Discovery in the Visual System, supported by Nati
                     Institute (NEI).
                     TAG_TISSUE=human optic nerve
                     TAG_LIB=UI-E-CQ1
                     TAG_SEQ=CCATTAAGTG"

ORIGIN
     Query Match   23.9%;   Score 328; DB 12;   Length 367;
     Best Local Similarity   100.0%; Pred. No. 1.1e-155;
     Matches 328;   Conservative   0; Mismatches   0; Indels   0; G:

Qy   1046 ACTCCCCCACCGCCGCCACTCTCCACCTCACTAGCTCCCCAATCCTGACCCTTTGAGG
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   367 ACTCCCCCACCGCCGCCACTCTCCACCTCACTAGCTCCCCAATCCTGACCCTTTGAGG

Qy   1106 CCAGTGATCTCGACTCCCCCTGGCCACAGACCCCCACTTCAGGCACTAAGAGGGGCTGGACCTA
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   307 CCAGTGATCTCGACTCCCCCTGGCCACAGACCCCCACTTCAGGCACTAAGAGGGGCTGGACCTA

Qy   1166 GTGGGCAAGGATGGGTCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTC
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   247 GTGGGCAAGGATGGGTCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTC

Qy   1226 GCAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAG
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   187 GCAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAG

Db   226 TGGGCAAGGATGGGTCCAGAAGACCCCACTCCACCTTCAGGCACTAAGGAGGGGCTGGACCTG

Qy   1227 CAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGA
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   166 CAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGA

Qy   1287 AGACAAGCTCCTCCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTATTTT
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   106 AGACAAGCTCCTCCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTATTTT

Qy   1347 TTATTGTGACAAAATGTTGATAAATGG 1373
             ||||||||||||||||||||||||||||
Db   46 TTATTGTGACAAAATGTTGATAAATGG 20

RESULT 30
AW195034/c
LOCUS   AW195034   337 bp   mRNA   linear   EST 2Ç
DEFINITION xn45g12.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2É
            mRNA sequence.
ACCESSION   AW195034
VERSION   AW195034.1   GI:6474026
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 337)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Mic
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequenciï
            Clone distribution: NCI-CGAP clone distribution informati
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq_primer: -40UP from Gibco
            High quality sequence stop: 330.
FEATURES   Location/Qualifiers
source   1..337
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:269671O"
            /lab_host="DH10B"
            /clone_lib="NCI_CGAP_Kid11"
            /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmaç
            a modified polylinker; Site_1: Not I; Site_2: Ecc
            Plasmid DNA from the normalized library NCI_CGAP
            prepared, and ss circles were made in vitro. Fol
            purification, this DNA was used as tracer in a s
            hybridization reaction. The driver was PCR-amplit
            from a pool of 5,000 clones made from the same li
            (cloneIDs 1322376-1323911, 1456007-1456775, and
            1500552-1502855). Subtraction by Bento Soares anc
            Fatima Bonaldo. "
ORIGIN

Query Match   23.7%;   Score 326;   DB 10;   Length 337;
Best Local Similarity   100.0%;   Pred. No. 1.2e-154;
Matches 326;   Conservative   0;   Mismatches   0;   Indels   0;   (

Qy   1048 TCCCCAACCGCCCACTCTCCACCTCACTAGCTCCCAATCCCTGACCCTTTGAGG(
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   337 TCCCCAACCGCCCACTCTCCACCTCACTAGCTCCCAATCCCTGACCCTTTGAGG(

CAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTAAAACAGATATTATTTTTATT 1345
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTTATTTTTATT 68

TTGTGACAAAATGTTGATAAATGG 1373
||||||||||||||||||||||||
TTGTGACAAAATGTTGATAAATGG 40

22
4.x1 NCI_CGAP_Brn23   346 bp   mRNA   linear   EST 29-SEP-2000
Homo sapiens cDNA clone IMAGE:3308766 3',
equence.
22
22.1   GI:10374253

apiens (human)
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 346)
NDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
al Cancer Institute / National Institute of Neurological
ers and Stroke, Brain Tumor Genome Anatomy Project
BTGAP), Tumor Gene Index
ished (1998)
t: Robert Strausberg, Ph.D.
cgapbs-r@mail.nih.gov
Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
o, Ph.D.
Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Library Arrayed by: Greg Lennon, Ph.D.
equencing by: Washington University Genome Sequencing Center
distribution: NCI-CGAP clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL, send email to:
mage.llnl.gov
imer: -40UP from Gibco
uality sequence stop: 344.
Location/Qualifiers
1..346
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3308766"
/tissue_type="glioblastoma (pooled)"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Brn23"
/note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGGAGCGGCCGCGAATACTTTTTTTTTTTTTTTTTTT
T 3'], double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldo."

arity   23.8%;   Score 327;   DB 10;   Length 346;
100.0%;   Pred. No. 3.6e-155;
onservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

CCCACCGCCCACTCTCCACCTCACTAGCTCCCAATCCCTGACCCTTTGAGGGCCCC 1106
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCACCGCCCACTCTCCACCTCACTAGCTCCCAATCCCTGACCCTTTGAGGGCCCC 287

TGATCTCGACTCCCCCTGGCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTG 1166
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGATCTCGACTCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTG 227

GCAAGGATGGGTCCAGAAGACCCCACTTCAGGCACTAAGGAGGGGCTGGACCTGGGGG 1226

```
Query Match          23.7%;  Score 326;  DB 10;  Length 345;
Best Local Similarity   100.0%;  Pred. No. 1.2e-154;
Matches  326;  Conservative   0;  Mismatches   0;  Indels   0;  Ga

Qy   1048 TCCCCACGCGCCACTCTCCACCTCACTAGCTCCCCAATCCTGACCCTTGAGGCC 1167
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    345 TCCCCACGCGCCACTCTCCACCTCACTAGCTCCCCAATCCTGACCCTTGAGGCC 218

Qy   1108 AGTGATCTGACTCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    285 AGTGATCTGACTCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTC

Qy   1108 AGTGATCTGACTCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTC 1287
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    285 AGTGATCTGACTCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTC 98

Qy   1168 GGGCAAGGATGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    225 GGGCAAGGATGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGC

Qy   1168 GGGCAAGGATGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGC 1347
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    225 GGGCAAGGATGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGC 38

Qy   1228 AGGAAGCCAAAGAGACTGGGCTAGGCCAGGAGTTCCCAAATGTGAGGGGCCAGAAA
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    165 AGGAAGCCAAAGAGACTGGGCTAGGCCAGGAGTTCCCAAATGTGAGGGGCCAGAAA

Qy   1288 GACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATATTATTTTTAT
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105 GACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATATTATTTTTAT

Qy   1348 TATTGTGACAAAATGTTGATAAATGG 1373
          ||||||||||||||||||||||||||
Db     45 TATTGTGACAAAATGTTGATAAATGG 20
```

```
RESULT 32
AW204512/c
LOCUS       AW204512              352 bp    mRNA      linear    EST 02-
DEFINITION  UI-H-BI1-aei-f-08-0-UI.s1 NCI_CGAP_Sub3 Homo sapiens cDNA c
            IMAGE:2719622 3', mRNA sequence.
ACCESSION   AW204512
VERSION     AW204512.1  GI:6503984
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelec
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 352)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project ((
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            The sequence contained an oligo-dT track that was present :
            oligonucleotide that was used to prime the synthesis of fi
            strand cDNA and therefore this may represent a bonafide po-
            tail  cDNA Library Preparation: M.B. Soares Lab Clone disti
            NCI-CGAP clone distribution information can be found throuc
            I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html The following repeti
            elements were found in this cDNA sequence: 49-81,
            >AT_rich#Low_complexity
            Seq_primer: M13 Forward
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1..352
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2719622"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_Sub3"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modif
                     polylinker; Site 1: Not I; Site 2: Eco RI; The
                     NCI_CGAP_Sub3 library is a subtracted library der:
                     the NCI_CGAP_Sub1 library, which is a subtracted
                     derived from BI.  BI constitutes a mixture of 21
                     normalized or subtracted NCI_CGAP libraries:
```

```
CTCGACTCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGT 1167
|||||||||||||||||||||||||||||||||||||||||||||||||||||
CTCGACTCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGT 218

AGGATGGGTCCAGAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGC 1227
||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGGATGGGTCCAGAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGC 158

GCCAAAGAGACTGGGCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAA 1287
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCCAAAGAGACTGGGCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAA 98

GCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTATTTTTATTAT 1347
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTATTTTTATTAT 38

TGACAAAATGTTGATAAATGG 1373
|||||||||||||||||||||
TGACAAAATGTTGATAAATGG 12
```

```
3
4.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:3406302 3'
  to contains element MSR1 repetitive element ;, mRNA
e.
3
3.1  GI:11452510
piens (human)
piens
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 345)
DS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
l Cancer Institute / National Institute of Neurological
rs and Stroke, Brain Tumor Genome Anatomy Project
TGAP), Tumor Gene Index
shed (1998)
: Robert Strausberg, Ph.D.
cgapbs-r@mail.nih.gov
Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,

ibrary Preparation: M. Bento Soares, Ph.D., M. Fatima
, Ph.D.
ibrary Arrayed by: Greg Lennon, Ph.D.
quencing by: Washington University Genome Sequencing Center
distribution: NCI-CGAP clone distribution information can be
hrough the I.M.A.G.E. Consortium/LLNL, send email to:
age.llnl.gov
mer: -40UP from Gibco
ality sequence stop: 333.
        Location/Qualifiers
        1..345
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:3406302"
        /lab_host="DH10B"
        /clone_lib="NCI_CGAP_Brn23"
        /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
        modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
        strand cDNA was primed with a Not I - oligo(dT) primer [5'
        TGTTACCAATCTGAAGTGGAGGGCGCCATATCTTTTTTTTTTTTTTTTTTTTTTTTTTT
        T 3'], double-stranded cDNA was ligated to Eco RI
        adaptors (Pharmacia),  digested with Not I and cloned into
        the Not I and Eco RI sites of  the modified pT7T3 vector.
        Library is normalized, and was constructed by Bento
        Soares and M.Fatima Bonaldo."
```

REFERENCE   1 (bases 1 to 364)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (
             Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             The sequence contained an oligo-dT track that was present
             oligonucleotide that was used to prime the synthesis of fi
             strand cDNA and therefore this may represent a bonafide po
             tail. cDNA Library Preparation: M.B. Soares Lab Clone dist
             NCI-CGAP clone distribution information can be found throu
             I.M.A.G.E. Consortium/LLNL at:
             www-bio.llnl.gov/bbrp/image/image.html The following repet
             elements were found in this cDNA sequence: 61-93,
             >AT_rich#Low complexity
             Seq primer: M13 Forward
             POLYA=Yes.

FEATURES   Location/Qualifiers
     source   1..364
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2724801"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_Sub4"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modi
                     polylinker; Site_1: Not I; Site_2: Eco RI; The
                     NCI_CGAP_Sub4 library is a subtracted library der
                     the NCI_CGAP_Sub2 library which is a subtracted
                     derived from the NCI_CGAP_Sub1 library, which is
                     subtracted library derived from BI.BI constitut
                     mixture of 21 normalized or subtracted NCI_CGAP
                     libraries: NCI_CGAP_Co4, NCI_CGAP_Pr22, NCI_CGAP_
                     NCI_CGAP_Co10, NCI_CGAP_Co16, NCI_CGAP_Kid5,
                     NCI_CGAP_Kid12, NCI_CGAP_Kid3, NCI_CGAP_Kid1,
                     NCI_CGAP_Lym2, NCI_CGAP_Br2, NCI_CGAP_Co8, NCI_C
                     NCI_CGAP_Lei2, NCI_CGAP_Brn23, NCI_CGAP_Lu5,
                     NCI_CGAP_Lu24, NCI_CGAP_Lu19, NCI_CGAP_GC4, NCI_C
                     NCI_CGAP_Brn25. These 21 libraries were pooled an
                     single-stranded DNA preparation of the resulting
                     was used as a tracer in a subtractive hybridizat
                     a driver whose composition is detailed below:
                     NCI_CGAP_Kid3 pool 1 : LLAM 3334-3337, 3682-3683,
                     3798-3803 (IMAGE CloneIDs 1322376-1323911,
                     1456008-1456775, 1500552-1502855) NCI_CGAP_Kid5 F
                     LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE Clon
                     1323912-1325831, 1471368-1472903, 1492104-1493255
                     NCI_CGAP_Lu5 pool 1 : LLAM 3575-3582, 3851-3854 (
                     CloneIDs 1414920-1417991, 1520904-1522439) NCI_CG
                     pool 1 : LLAM 3164-3167, 3716-3720, 3733-3735 (IM
                     CloneIDs 1257096-1258631, 1469064-1470983,
                     1475592-1476743) NCI_CGAP_Pr22 pool 1 : LLAM 2457
                     2758-2759, 3062-3068 (IMAGE CloneIDs 985608-98675
                     1101192-1101959, 1217928-1220615) NCI_CGAP_Co10 F
                     LLAM 2644-2653, 2871-2872 (IMAGE CloneIDs 1057416
                     1144584-1145351) Subtraction was performed as pre
                     described [Bonaldo, Lennon & Soares (1996): Norm
                     and Subtraction: Two Approaches To Facilitate Ge
                     Discovery. Genome Research 6, 791-806.]
                     TAG_TISSUE=kidney
                     TAG_LIB=NCI_CGAP_Kid5
                     TAG_SEQ=ATTC"

ORIGIN

             Query Match   23.7%; Score 326; DB 10; Length 364;
             Best Local Similarity 100.0%; Pred. No. 1.2e-154;
             Matches 326; Conservative   0; Mismatches   0; Indels   0; (

Qy   1048 TCCCCCACCGCCCACTCTCCACCTCACTAGCTCCCCAATCCCTGACCCTTTGAGG(
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   364 TCCCCCACCGCCCACTCTCCACCTCACTAGCTCCCCAATCCCTGACCCTTTGAGG(

NCI_CGAP_Co4, NCI_CGAP_Pr22, NCI_CGAP_Pr28, NCI_CGAP_Co10,
NCI_CGAP_Co16, NCI_CGAP_Kid5, NCI_CGAP_Kid12,
NCI_CGAP_Kid3, NCI_CGAP_Kid1, NCI_CGAP_Lym2, NCI_CGAP_Lei2,
NCI_CGAP_Br2, NCI_CGAP_Co8, NCI_CGAP_Li1, NCI_CGAP_Lu24,
NCI_CGAP_Brn23, NCI_CGAP_Lu5, NCI_CGAP_GC6,
NCI_CGAP_Lu19, NCI_CGAP_GC4, NCI_CGAP_GC4,
NCI_CGAP_Brn25. These 21 libraries were pooled and a
single-stranded DNA preparation of the resulting mixture
was used as a tracer in a subtractive hybridization with
a driver whose composition is detailed below:
NCI_CGAP_Kid3 pool 1   LLAM 3334-3337, 3682-3683,
3798-3803 (IMAGE CloneIDs 1322376-1323911,
1456008-1456775, 1500552-1502855); NCI_CGAP_Kid5 pool 1
LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE CloneIDs
1323912-1325831, 1471368-1472903, 1492104-1493255);
NCI_CGAP_Lu5 pool 1   LLAM 3575-3582, 3851-3854 (IMAGE
CloneIDs 1414920-1417991, 1520904-1522439); NCI_CGAP_GC4
pool 1   LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE
CloneIDs 1257096-1258631, 1469064-1470983,
1475592-1476743); NCI_CGAP_Pr22 pool 1   LLAM 2457-2459,
2758-2759, 3062-3068 (IMAGE CloneIDs 985608-986759,
1101192-1101959, 1217928-1220615); NCI_CGAP_Co10 pool 1
LLAM 2644-2653, 2871-2872 (IMAGE CloneIDs 1057416-1061255,
1144584-1145351). Subtraction was performed as normalization
described [Bonaldo, Lennon & Soares (1996): Normalization
and Subtraction: Two Approaches To Facilitate Gene
Discovery. Genome Research 6, 791-806.
TAG_TISSUE=breast
TAG_LIB=NCI_CGAP_Br2
TAG_SEQ=AAACC"

arity   23.7%; Score 326; DB 10; Length 352;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

CCACCGCCCACTCTCCACCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCC 1107
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCACCGCCCACTCTCCACCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCC 293

ATCTCGACTCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGT 1167
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ATCTCGACTCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGT 233

AAGGATGGGTCCAGAAGACCCCCACTTCAGGCACTAAGAGGGGACCTGGCGGC   1227
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AAGGATGGGTCCAGAAGACCCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGC 173

AGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCCAGAAACAA 1287
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCCAGAAACAA 113

AGCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTATTTTTATTAT 1347
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTATTTTTATTAT   53

GTGACAAAATGTTGATAAATGG 1373
||||||||||||||||||||||
GTGACAAAATGTTGATAAATGG   27

520
BI2-agn-b-05-0-UI.s1 NCI_CGAP_Sub4 Homo sapiens cDNA clone
2724801 3', mRNA sequence.
520
520.1 GI:6698256

sapiens (human)
sapiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
lia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

364 bp   mRNA   linear   EST 16-JAN-2000

```
Qy   1050  CCCCACCGCCCACTCTCCCACCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCC
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    347  CCCCACCGCCCACTCTCCCACCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCC

Qy   1110  TGATCTCGACTCCCCCGTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTG
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    287  TGATCTCGACTCCCCCGTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTG

Qy   1170  GCAAGGATGGGTCCAGAAGACACCCACTTCAGCACTAAGAGGGGCTGGACCTGGCGG
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    227  GCAAGGATGGGTCCAGAAGACACCCACTTCAGCACTAAGAGGGGCTGGACCTGGCGG

Qy   1230  GAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACA
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    167  GAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACA

Qy   1290  CAAGCTCCTCCCTTGAGAATTCCTGTGGGATTTTTAAAACAGATATTATTTTTATTA
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107  CAAGCTCCTCCCTTGAGAATTCCTGTGGGATTTTTAAAACAGATATTATTTTTATTA

Qy   1350  TTGTGACAAAATGTTGATAAATGG  1373
                ||||||||||||||||||||||||
Db     47  TTGTGACAAAATGTTGATAAATGG    24
```

RESULT 35
BI966255
LOCUS   BI966255   456 bp   mRNA   linear   EST 12-
DEFINITION   ie72g04.y1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sa
             cDNA clone IMAGE:5672623 5' similar to TR:O43508 O43508 TNF
             WEAK INDUCER OF APOPTOSIS ;, mRNA sequence.
ACCESSION   BI966255
VERSION   BI966255.1   GI:16340660
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelec
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 456)
AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner
             Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton
             Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Bistai
             Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
             Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili
             Williams,T., Jackson,Y. and Bowers,Y.
             Endocrine Pancreas Consortium
TITLE   Unpublished (2000)
JOURNAL   Other ESTs: ie72g04.x1
COMMENT   Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
             Endocrine Pancreas Consortium
             Harvard University, Howard Hughes Medical Institute
             Dept of Molecular and Cellular Biology, 7 Divinity Ave, Can
             MA 02138
             Tel: 617-495-1812
             Fax: 617-495-8557
             Email: dmelton@biohp.harvard.edu
             Library was constructed by Dr. Douglas Melton DNA sequencin
             Washington University Genome Sequencing Center For informat
             obtaining a clone please contact: Juliana Brown
             (brown@fas.harvard.edu) This sequence now available from th
             consortium, for clone orders contact: info@image.llnl.gov
             High quality sequence stop: 429.
FEATURES   Location/Qualifiers
     source   1..456
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5672623"
                     /sex="Both"
                     /tissue_type="Islets of Langerhans"
                     /dev_stage="Adult"
                     /lab_host="DH10B"
                     /clone_lib="Melton Normalized Human Islet 4 N4-HIS

```
Qy   1167  CTCGACTCCCCCGTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGT  1167
                |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    245  CTCGACTCCCCCGTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGT   245

Qy   1227  AGGATGGGTCCAGAAGACACCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGC  1227
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    185  AGGATGGGTCCAGAAGACACCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGC   185

Qy   1287  GCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAA  1287
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    125  GCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAA   125

Qy   1347  CTCCTCCCTTGAGAATTCCCTGTGGGATTTTTAAAACAGATATTTTTATTAT  1347
                ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     65  CTCCTCCCTTGAGAATTCCCTGTGGGATTTTTAAAACAGATATTTTTATTAT    65

Qy   1373  TGACAAAATGTTGATAAATGG  1373
                |||||||||||||||||||||
Db     39  TGACAAAATGTTGATAAATGG    39
```

             399 bp   mRNA   linear   EST 28-JUL-1999
.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2297856 3',
quence.
1
1.1   GI:5633396
piens (human)
piens
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 399)
P http://www.ncbi.nlm.nih.gov/ncicgap.
1 Cancer Institute, Cancer Genome Anatomy Project (CGAP),
ene Index
shed (1997)
: Robert Strausberg, Ph.D.
cgapbs-r@mail.nih.gov
Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Buck, M.D., Ph.D.
ibrary Preparation: M. Bento Soares, Ph.D.
/clone lib="NCI CGAP Kid1"
quencing Arrayed by: Greg Lennon, Ph.D.
quencing by: Washington University Genome Sequencing Center
distribution: NCI-CGAP clone distribution information can be
hrough the I.M.A.G.E. Consortium/LLNL at:
.llnl.gov/bbrp/image/image.html
mer: -40UP from Gibco.
Location/Qualifiers
1..399
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2297856"
                     /lab_host="DH10B"
                     /clone lib="NCI CGAP Kid1"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site 1: Not I; Site 2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP Kid3 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIds 1322376-1323911, 1456007-1456775, and
                     1500552-1502855). Subtraction by Bento Soares and M.
                     Fatima Bonaldo. "

rity   23.6%;   Score 324;   DB 9;   Length 399;
rity   100.0%;   Pred. No. 1.3e-153;
nservative   0; Mismatches   0; Indels   0; Gaps   0;

University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 3054 row: O column: 5
Class: BAC ends
High quality sequence stop: 367.

```
FEATURES             Location/Qualifiers
     source          1..367
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=3054 Col=5 Row=O"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Libr
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clon
                     E-Coli DH10B"
ORIGIN
```

/note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;
Site 2: Sal 1; Starting library constructed using
SuperScript Plasmid Library kit (Life Technologies). cDNA
made by oligo-dT priming. Size-selected by column
fractionation; average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an Rcot of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."

```
Query Match                  23.4%;    Score 321;   DB 28;  Length 367;
Best Local Similarity  100.0%;  Pred. No. 4.2e-152;  Indels    0;   G
Matches  321; Conservative   0; Mismatches   0; Indels    0;

arity   23.6%;    Score 324;  DB 12;   Length 456;
Pred. No. 1.3e-153;    Indels    0;  Gaps    0;
onservative   0;  Mismatches   1;  Indels    0;  Gaps    0;

Qy   776 TGCGGATCCGACCCTCCCTGGGCCATCTCAAGGGTCGCCCCTTCTCACCTAC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TGCGGATCCGACCCTCCCTGGGCCATCTCAAGGGTCGCCCCTTCTCACCTAC

CCCTCGGGCCCCAGCTCCGCCTCCGCCAGGTGTCTGGGCTGTTGGCCCTGCGGCCA  765
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCTCGGGCCCCAGCTCCGCCTCCGCCAGGTGTCTGGGCTGTTGGCCCTGCGGCCA  141

Qy   836 GACTCTTCCAGGTTCACTGAGGGGCCCTGGTCTCCCCACAGTCGTCCCAGGCTGCC
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 GACTCTTCCAGGTTCACTGAGGGGCCCTGGTCTCCCCACAGTCGTCCCAGGCTGCC

CCTCCCTGCGGATCGGACCCTCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTC  825
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCTCCCTGCGGATCGGACCCTCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTC  201

ACTTCGGACTCTTCCAGGTTCACTGAGGGGCCCTGGTCTCCCCACAGTCGTCCCAG  885
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
ACTTCGGACTCTTCCAGGTTCACTGAGGGGCCCTGGTCTCCCCACAGTCGTCCCAG  261

Qy   896 CCCCTCGACAGCTCTCTGGGCACCCGGTCCCCTGCCCCACCCTCAGCGCGTCT
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CCCCTCGACAGCTCTCTGGGCACCCGGTCCCCTGCCCCACCCTCAGCGCGTCT

CCGGCTCCCCTGACAGCTCTCTGGGCACCCGGTCCCCTGCCCCACCCTCAGCC  945
|||||||||||||||||||||||||||||||||||||||||||||||||||||
CCGGCTCCCCTGACAGCTCTCTGGGCACCCGGTCCCCTGCCCCACCCTCAGCC  321

Qy   956 CCAGACCTGCCCCTCCCTCTAGAGGCTGCCTGTTCACGTGTTTTCATCC
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCAGACCTGCCCCTCCCTCTAGAGGCTGCCTGTTCACGTGTTTTCATCC

TTTGCTCCAGACCTGCCCCTCCCTCTAGAGGCTGCCTGTTCACGTGTTTT  1005
||||||||||||||||||||||||||||||||||||||||||||||||||
TTTGCTCCAGACCTGCCCCTCCCTCTAGAGGCTGCCTGTTCACGTGTTTT  381

Qy  1016 TAAATACAGTATTCCCACTCTTATCTTACAACTCCCCCACCGCCCACTCTCCACCT
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TAAATACAGTATTCCCACTCTTATCTTACAACTCCCCCACCGCCCACTCTCCACCT

CCCACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCCACCGCCCACTCT  1065
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCCACCGCCCACTCT  441

Qy  1076 AGCTCCCAATCCCTGACCCT  1096
         ||||||||||||||||||||
Db   301 AGCTCCCAATCCCTGACCCT  321

CTCACTAGCTC  1080
|||||||||||
CTCACTAGCTC  456
```

65   367 bp   DNA   linear   GSS 27-AUG-1998
4_A1_H03_MF CIT Approved Human Genomic Sperm Library D Homo
s genomic clone Plate=3054 Col=5 Row=O, genomic survey
ce.
65
65.1 GI:3471394

apiens (human)
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 367)
:as,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
:,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
ce-tagged connectors: A sequence approach to mapping and
ing the human genome
Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
589
764
ct: Mahairas GG, Wallace JC, Hood L
Throughput Sequencing Center

Left column

nrough the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov
LLAM12814 row: n column: 16
ality sequence start: 74
ality sequence stop: 420.
Location/Qualifiers
1..1819
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5763279"
/lab_host="DH10B"
/clone_lib="NIH_MGC_114"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."

```
rity     23.4%;   Score 321;  DB 12;  Length 1819;
100.0%;  Pred. No. 5.9e-152;
onservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

CCACTCTCCACCTCACTAGCTCCCCAATCCCTGACCCTTGAGGCCCCCAGTGA 1112
|||||||||||||||||||||||||||||||||||||||||||||||||||||
CCACTCTCCACCTCACTAGCTCCCCAATCCCTGACCCTTGAGGCCCCCAGTGA 139

ACTCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCA 1172
||||||||||||||||||||||||||||||||||||||||||||||||||||||
ACTCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCA 199

GGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGCCAGGAA 1232
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGCCAGGAA 259

AGAGACTGGGCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACA 1292
||||||||||||||||||||||||||||||||||||||||||||||||||||
AGAGACTGGGCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACA 319

TCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTATTTTTATTATTATTG 1352
||||||||||||||||||||||||||||||||||||||||||||||||||||
TCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTATTTTTATTATTATTG 379

AAAATGTTGATAAATGG 1373
|||||||||||||||||
AAAATGTTGATAAATGG 400
```

19   374 bp   mRNA   linear   EST 14-FEB-2002
5.x1 Human insulinoma Homo sapiens cDNA 3', mRNA sequence.
19
19.1 GI:18669065

apiens (human)
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 374)
D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
ka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
r,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
t.A., Theising,B., Ritter,R., Ronko,I., Bennett,J.,
as,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
ms,T., Jackson,Y. and Bowers,Y.
ine Pancreas Consortium
ished (2000)

Right column

COMMENT   Other_ESTs: ih15b05.y1
          Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
          Endocrine Pancreas Consortium
          Harvard University, Howard Hughes Medical Institute
          Dept of Molecular and Cellular Biology, 7 Divinity Ave, Can
          MA 02138
          Tel: 617-495-1812
          Fax: 617-495-8557
          Email: dmelton@biohp.harvard.edu
          Library was constructed by Dr. J. Ferrer In vivo mass-excis
          pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washingt
          University Genome Sequencing Center For information on obta
          clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.ec
          Seq primer: -40UP from Gibco.

FEATURES Location/Qualifiers
source   1..374
          /organism="Homo sapiens"
          /mol_type="mRNA"
          /db_xref="taxon:9606"
          /tissue_type="insulinoma"
          /lab_host="DH10B (phage-resistant)"
          /clone_lib="Human insulinoma"
          /note="Organ: pancreas; Vector: pBluescript SK-;
          XhoI; Site 2: EcoRI; Constructed with lambda ZAPI
          (Stratagene) by Dr. J. Ferrer, in vivo mass-excise
          pBluescript SK- by Dr. H. Inoue following the Was
          University protocol
          (http://genome.wustl.edu/est/lambda_protocol.shtml
          Please contact Hiroshi Inoue, MD/PhD for further
          information on this library (Metabolism Division
          Laboratory, Washington University School of Medic
          8127, 660 S Euclid Ave, St. Louis, MO 63110). Not
          is a Washington University Pancreas EST project 1

ORIGIN

```
Query Match       23.3%;  Score 320;  DB 12;  Length 374;
Best Local Similarity 100.0%;  Pred. No. 1.4e-151;
Matches 320;  Conservative 0;  Mismatches 0;  Indels 0;  G

Qy  1054 ACGCGCCACTCTCCACCTCACTAGCTCCCCAATCCCTGACCCTTGAGGCCCCCAG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   346 ACGCGCCACTCTCCACCTCACTAGCTCCCCAATCCCTGACCCTTGAGGCCCCCAG

Qy  1114 CTCGACTCCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   286 CTCGACTCCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGG

Qy  1174 GGATGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGGCAG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   226 GGATGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGGCAG

Qy  1234 CCAAAGAGACTGGGCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAACAAGA
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   166 CCAAAGAGACTGGGCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAACAAGA

Qy  1294 CTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTTTTTATTATTA
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   106 CTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTTTTTATTATTA

Qy  1354 GACAAAATGTTGATAAATGG 1373
         ||||||||||||||||||||
Db    46 GACAAAATGTTGATAAATGG 27
```

RESULT 39
AI695776/c
LOCUS   AI695776   329 bp   mRNA   linear   EST 17
DEFINITION wb77g07.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:231
            mRNA sequence.
ACCESSION   AI695776
VERSION   AI695776.1 GI:4983676
KEYWORDS   EST.
SOURCE   Homo sapiens (human)

```
LOCUS       BF195436                340 bp    mRNA    linear   EST 03
DEFINITION  7n17g12.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:35
            mRNA sequence.
ACCESSION   BF195436
VERSION     BF195436.1  GI:11082306
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 340)
AUTHORS     NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute / National Institute of Neurolog
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGAP), Tumor Gene Index
JOURNAL     Unpublished (1998)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfe
            Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fati
            Bonaldo, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencin
            Clone distribution: NCI-CGAP clone distribution informati
            found through the I.M.A.G.E. Consortium/LLNL, send email t
            info@image.llnl.gov.
FEATURES             Location/Qualifiers
     source          1..340
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3564887"
                     /tissue_type="glioblastoma (pooled)"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Brn23"
                     /note="Organ: brain; Vector: pT7T3D-Pac (Pharmac
                     modified polylinker; Site 1: Not I; Site 2: Eco I
                     strand cDNA was primed with a Not I - oligo(dT) I
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATATCTTTTTTTTTTTT
                     T 3'); double-stranded cDNA was ligated to Eco I
                     adaptors (Pharmacia), digested with Not I and c
                     the Not I and Eco RI sites of  the modified pT7T
                     Library is normalized, and was constructed  by B
                     Soares and M.Fatima Bonaldo."
ORIGIN
            Query Match         23.2%;   Score 318;  DB 10;  Length 340;
            Best Local Similarity  100.0%;   Pred. No. 1.4e-150;  Indels     0;
            Matches   318;  Conservative    0;  Mismatches    0;  Indels    0;

       1056 CGCCCACTCTCACCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCCAGT
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
        340 CGCCCACTCTCCACCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCCAGT

       1116 CGACTCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGG
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
        280 CGACTCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGG

       1176 ATGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGCAGG
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
        220 ATGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGCAGG

       1236 AAAGAGACTGGGCTTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGAC
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
        160 AAAGAGACTGGGCTTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGAC

       1296 CCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTTTTATTATTAT
            |||||||||||||||||||||||||||||||||||||||||||||
        100 CCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTTTATTTATTAT

       1356 CAAAATGTTGATAAATGG 1373
            ||||||||||||||||||
```

```
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 329)
AP http://www.ncbi.nlm.nih.gov/ncicgap.
al Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Gene Index
ished (1997)
t: Robert Strausberg, Ph.D.
    cgapbs-r@mail.nih.gov
    Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
-Buck, M.D., Ph.D.
    Library Preparation: M. Bento Soares, Ph.D.
    Library Arrayed by: Greg Lennon, Ph.D.
    equencing by: Washington University Genome Sequencing Center
    distribution: NCI-CGAP clone distribution information can be
    through the I.M.A.G.E. Consortium/LLNL at:
    o.llnl.gov/bbrp/image.html
    Length: 305   Std Error: 0.00
imer: -40UP from Gibco
uality sequence stop: 302.
            Location/Qualifiers
            1..329
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:2311740"
            /sex="male"
            /dev_stage="adult"
            /lab_host="DH10B"
            /clone_lib="NCI_CGAP_Pr28"
            /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
            with a modified polylinker; Plasmid DNA from the
            normalized library NCI_CGAP_Pr22 was prepared, and ss
            circles were made in vitro. Following HAP purification,
            this DNA was used as tracer in a subtractive hybridization
            reaction. The driver was PCR-amplified cDNAs from a pool
            of 5,000 clones made from the same library (cloneIDs
            985608-986759, 1101192-1101959, and 1217928-1220615).
            Subtraction by Bento Soares and M. Fatima Bonaldo."

            23.2%;   Score 318;  DB 9;  Length 329;
arity   100.0%;   Pred. No. 1.4e-150;  Indels     0;   Gaps     0;
onservative    0;  Mismatches    0;  Indels    0;
```

```
CACTCTCCACCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCAGTGATCT 1115
|||||||||||||||||||||||||||||||||||||||||||||||||||||
CACTCTCCACCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCAGTGATCT 270

TCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGG 1175
||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGG 210

GGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGCGGAGGAGCC 1235
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGCGGAGGAGCC 150

GAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCT 1295
|||||||||||||||||||||||||||||||||||||||||||||||||||||
GAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCT 90

CCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTTTATTTATTATTGTGA 1355
|||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTTTATTTATTATTGTGA 30

AATGTTGATAAATGG 1373
|||||||||||||||
AATGTTGATAAATGG 12
```

Qy   1358   AAATGTTGATAAATGG   1373
            |||||||||||||||
Db   35   AAATGTTGATAAATGG   20

RESULT 42
AI291866/c
LOCUS   AI291866
DEFINITION   qm86c02.x1 NCI_CGAP_Lu5 Homo sapiens cDNA clone IMAGE:18956
            mRNA sequence.
ACCESSION   AI291866
VERSION   AI291866.1   GI:3934640
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelec
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 416)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project ((
            Tumor Gene Index
  TITLE   Unpublished (1997)
  JOURNAL
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Mich
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing
            Clone distribution: NCI-CGAP clone distribution informati
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 649   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 411.
FEATURES   Location/Qualifiers
  source   1..416
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:1895618"
            /tissue_type="carcinoid"
            /lab_host="DH10B"
            /clone_lib="NCI_CGAP_Lu5"
            /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia
            modified polylinker; 1st strand cDNA was prepared
            neuroendocrine lung carcinoid, and was then prime
            Not I - oligo(dT) primer. Double-stranded cDNA wa
            to Eco RI adaptors (Pharmacia), digested with Not
            cloned into the Not I and Eco RI sites of the mod
            pT7T3 vector. Library is normalized. Library was
            constructed by Bento Soares and M. Fatima Bonaldo

ORIGIN
Query Match   22.5%;   Score 309;   DB 9;   Length 416;
Best Local Similarity   100.0%;   Pred. No. 5.7e-146;
Matches 309;   Conservative   0;   Mismatches   0;   Indels   0;   G

Qy   1065   TCCACCTCACTAGCTCCCCAATCCCTGACCCTTGAGGCCCCAGTGATCTCGACT
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   334   TCCACCTCACTAGCTCCCCAATCCCTGACCCTTGAGGCCCCAGTGATCTCGACT

Qy   1125   CCTGGCCACAGACCCCAGGGCATTGTGTTCACTGTACTCTGTGGCAAGGATGGG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   274   CCTGGCCACAGACCCCAGGGCATTGTGTTCACTGTACTCTGTGGCAAGGATGGG

Qy   1185   GAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGCGGAAGCCAAAGA
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   214   GAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGCGGAAGCCAAAGA

Qy   1245   GGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCTCC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   154   GGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCTCC

TGTTGATAAATGG 23

5
.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
842906 3', mRNA sequence.
5
5.1 GI:3804188

piens (human)
piens
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 407)
P http://www.ncbi.nlm.nih.gov/ncicgap.
l Cancer Institute, Cancer Genome Anatomy Project (CGAP),
ene Index
shed (1997)
: Robert Strausberg, Ph.D.
cgapbs-r@mail.nih.gov
one is available royalty-free through LLNL ; contact the
onsortium (info@image.llnl.gov) for further information.
Length: 610   Std Error: 0.00
mer: -40UP from Gibco
ality sequence stop: 401.
Location/Qualifiers
1..407
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1842906"
/lab_host="DH10B"
/clone_lib="Soares_NFL_T_GBC_S1"
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NbHL19W, testis NHT, and B-cell
NCI_CGAP_GCB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo. "

rity   23.0%;   Score 316;   DB 9;   Length 407;
100.0%;   Pred. No. 1.5e-149;
nservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

TCTCCACCTCACTAGCTCCCCAATCCCTGACCCTTGAGGCCCCAGTGATCTCG 1117
||||||||||||||||||||||||||||||||||||||||||||||||||||
CTCCACCTCACTAGCTCCCCAATCCCTGACCCTTGAGGCCCCAGTGATCTCG 276

CCCCTGGCCACAGACCCCAGGGCATTGTGTTCACTGTACTCTGTGGCAAGGAT 1177
||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCCTGGCCACAGACCCCAGGGCATTGTGTTCACTGTACTCTGTGGCAAGGAT 216

CCAGAAGACCCCACTTCAGGCACTAAGAGGGGCGGACCTGGCGGCGGAAGCCAA 1237
||||||||||||||||||||||||||||||||||||||||||||||||||||
CCAGAAGACCCCACTTCAGGCACTAAGAGGGGCGGACCTGGCGGCGGAAGCCAA 156

ACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCC 1297
||||||||||||||||||||||||||||||||||||||||||||||||||||
ACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCC 96

TTGAGAATTCCCTGTGGATTTTTAAAACAGATATTATTTTTATTATTATTGTGACA 1357
||||||||||||||||||||||||||||||||||||||||||||||||||||
TTGAGAATTCCCTGTGGATTTTTAAAACAGATATTATTTTTATTATTATTGTGACA 36

```
Qy    867  CTCCCCACAGTCGTCCCAGCTGCGGCTCCCCGACCTGGTCTCTGGGCACCCGG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240  CTCCCCACAGTCGTCCCAGCTGCGGCTCCCCGACCTGGTCTCTGGGCACCCGG  35

Qy    927  CTCTGCCCCACCCCAGCCCGCTCTTTGCTCCAGACCTGCCCCTCCCTCTAGAGGCT
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300  CTCTGCCCCACCCCAGCCCGCTCTTTGCTCCAGACCTGCCCCTCCCTCTAGAGGCT

Qy    987  GGGCCGTTCACGTGTTTCATCCACATAAATACAGTATTCCCACTCTTATCTT
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360  GGGCCGTTCACGTGTTTCATCCACATAAATACAGTATTCCCACTCTTATCTT

Qy   1047  CTCCCCA  1054
           |||||||
Db    420  CTCCCCA  427
```

RESULT 44

AI202121/c

LOCUS   AI202121   317 bp   mRNA   linear   EST 02
DEFINITION qi52c03.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:18
            mRNA sequence.
ACCESSION   AI202121
VERSION   AI202121.1 GI:3754727
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 317)
AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute / National Institute of Neurolog
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGAP), Tumor Gene Index
JOURNAL   Unpublished (1998)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfe
            Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fati
            Bonaldo, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencin
            Clone distribution: NCI-CGAP clone distribution informati
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1005   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 308.
FEATURES   Location/Qualifiers
     source   1..317
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1860100"
                     /tissue_type="anaplastic oligodendroglioma"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Brn25"
                     /note="Organ: brain; Vector: pT7T3D-Pac (Pharmac
                     modified polylinker; Site 1: Not I; Site 2: Eco I
                     strand cDNA was primed with a Not I - oligo(dT)]
                     TGTTACCAATCTGAAGTGGGAGCGGCGCCATAGGTTTTTTTTTTTTTT
                     T 3'l, double-stranded cDNA was ligated to Eco I
                     adaptors (Pharmacia), digested with Not I and c
                     the Not I and Eco RI sites of the modified pT7T
                     Library is normalized, and was constructed by B
                     Soares and M.Fatima Bonaldo."

ORIGIN

Query Match   22.1%;   Score 303; DB 9; Length 317;
Best Local Similarity 100.0%; Pred. No. 6.3e-143;
Matches 303; Conservative 0; Mismatches 0; Indels 0; '
```

Left column (rotated):

```
            TTCCCTGTGGATTTTTAAAAACAGATATTATTTTTATTATTATTGTGACAAAATGTT  1364
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
            TTCCCTGTGGATTTTTAAAACAGATATTATTTTTTATTATTATTGTGTGACAAAATGTT  35

            AATGG  1373
            |||||
            AATGG  26

63          667F1 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4367225 5',
            sequence.
63
63.1        GI:12603569
```

LOCUS definition block:
```
            910 bp    mRNA    linear    EST 30-JAN-2001
```

```
apiens (human)
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 910)
C http://mgc.nci.nih.gov/.
al Institutes of Health, Mammalian Gene Collection (MGC)
ished (1999)
t: Robert Strausberg, Ph.D.
cgapbs-r@mail.nih.gov
Procurement: ATCC
Library Preparation: Life Technologies, Inc.
Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
equencing by: Incyte Genomics, Inc.
distribution: MGC clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL at:
/image.llnl.gov
LLAM10019  row: e  column: 18
uality sequence stop: 493.
             1..910
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:4367225"
             /tissue_type="osteosarcoma, cell line"
             /lab_host="DH10B (phage-resistant)"
             /clone_lib="NIH_MGC_86"
             /note="Organ: bone; Vector: pCMV-SPORT6; Site 1: NotI;
             Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
             Average insert size 1.533 kb. Library enriched for
             full-length clones and constructed by Life Technologies.
             Note: this is a NIH_MGC Library."
```

```
arity           22.4%;   Score 307;  DB 12;  Length 910;
               99.8%;   Pred. No. 7.1e-145;
onservative  0;  Mismatches  0;  Indels  1;  Gaps  1;
```

```
CTACCTGAAGCTGGACTTGCTGGTGATGGTGTGCTGGCCCTGGCGCCTGCTGGAGGA  686
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TACCTGAAGCTGGACTTGCTGGTGGTGATGGTGTGCTGGCCCTGCGCTGCCTGGGAGGA  60

CTCAGCCACTGCGCGGCCAGTTCCCTGGGCCCCAGTCTCCGCCTCCTGCCAGTGTCTGG  746
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTCAGCCACTGCGCGGCCAGTTCCCTGGGCCCCAGCCCCAGTCTCCGCCTCCTGCCAGTGTCTGG  120

GTTGGGCCCTGCGGCCGGAGGTCCTCCCCTGCGGATCCCCGCACCCTCCCCTGGGCCCATCT  806
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTTGGGCCCTGCGGCCA-GGTCCTCCCCTGCGGATCCCCGCACCCTCCCCTGGGCCCATCT  179

GGCTGCCCCCTTCCTCACCTACTTCGGACTCTTCCAGGTTCACTGAGGGGGCCCTGGT  866
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGCTGCCCCCTTCCTCACCTACTTCGGACTCTTCCAGGTTCACTGAGGGGGCCCTGGT  239
```

following tissue(s): retina. The library was const
according to Bonaldo, Lennon and Soares. Genome Re
6:791-806, 1996. First strand cDNA synthesis was p
with an oligo-dT primer containing a Not I site. D
stranded cDNA was ligated to an EcoR I adaptor, di
with Not I, and cloned directionally into pT7T3-Pa
vector. The oligonucleotide used to prime the synt
first-strand cDNA contains a library tag sequence
located between the Not I site and the (dT)18 tail
sequence tag for this library is CGGCG. This libra
created for the program, Gene Discovery in the Vis
System, supported by National Eye Institute (NEI)

ORIGIN

Query Match   21.7%;   Score 298; DB 12; Length 346;
Best Local Similarity 100.0%;   Pred. No. 2.3e-140;
Matches 298;   Conservative 0; Mismatches 0; Indels 0; G

Qy 1076 AGCTCCCAATCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCCCCCCTGGCC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 AGCTCCCAATCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCCCCCCTGGCC

Qy 1136 ACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 ACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACC

Qy 1196 TTCAGGCACTAAGAGAGGGGCTGGACCTGGCGCGGCAGGAAGCCAAAGAGACTGGGCCTA
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TTCAGGCACTAAGAGAGGGGCTGGACCTGGCGCGGCAGGAAGCCAAAGAGACTGGGCCTA

Qy 1256 AGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTC

Qy 1316 TGGATTTTTAAAACAGATATTATTTTTATTATTATTTGTGACAAAATGTTGATAAAT
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TGGATTTTTAAAACAGATATTATTTTTATTATTATTTGTGACAAAATGTTGATAAAT

RESULT 46
BU729427/c
LOCUS   BU729427   333 bp   mRNA   linear   EST 09
DEFINITION UI-E-CL1-afe-m-20-0-UI s1 UI-E-CL1 Homo sapiens cDNA clone
            UI-E-CL1-afe-m-20-0-UI 3', mRNA sequence.
ACCESSION   BU729427
VERSION   BU729427.1 GI:23652299
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 333)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE   Normalization and subtraction: two approaches to facilitat
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
   MEDLINE   97044477
   PUBMED   8889548
COMMENT   Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road, 4156   MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
            cDNA Library preparation: Dr. M. Bento Soares, University
            cDNA Library Arrayed by: Dr. M. Bento Soares, University
            DNA Sequencing by: Dr. M. Bento Soares, University of Iow
            Clone Distribution: Researchers may obtain clones from Re
            Genetics (www.resgen.com).
            The following repetitive elements were found in this cDNA
            sequence: 60-92, >AT_rich#Low_complexity (matched complime

JOURNAL   Genome Res. 6 (9), 791-806 (1996)
MEDLINE   97044477
PUBMED   8889548
COMMENT   Contact: Soares, MB
           Coordinated Laboratory for Computational Genomics
           University of Iowa
           375 Newton Road   4156   MEBRF, Iowa City, IA 52242, USA
           Tel: 319 335 8250
           Fax: 319 335 9565
           Email: bento-soares@uiowa.edu
           Tissue Procurement: Dr. Gregg Hageman
           cDNA library preparation: Dr. M. Bento Soares, Univeristy
           cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy
           DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iow
           Clone Distribution: Researchers may obtain clones from Re
           Genetics (www.resgen.com).
           Seq primer: M13 Reverse.

FEATURES   Location/Qualifiers
     source   1..298
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="UI-E-CQ1-acs-a-05-0-UI"
                  /tissue_type="optic nerve"
                  /dev_stage="adult"
                  /lab_host="DH10B (Life Technologies) (T1 phage re
                  /clone_lib="UI-E-CQ1"
                  /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia)
                  modified polylinker; Site 1: EcoR I; Site 2: Not
                  UI-E-CQ1 is a normalized cDNA library containing
                  following tissue(s): optic nerve. The library was
                  constructed according to Bonaldo, Lennon and Soar
                  Genome Research, 6:791-806, 1996. First strand cD
                  synthesis was primed with an oligo-dT primer cont
                  Not I site. Double stranded cDNA was ligated to a
                  adaptor, digested with Not I, and cloned directio
                  into pT7T3-Pac vector. The oligonucleotide used t
                  the synthesis of first-strand cDNA contains a lib
                  sequence that is located between the Not I site a
                  (dT)18 tail. The sequence tag for this library is
                  CCATTAAGTG. This library was created for the prog
                  Discovery in the Visual System, supported by Nati
                  Institute (NEI)."

ORIGIN

Query Match   21.1%; Score 290; DB 12; Length 298;
Best Local Similarity   100.0%; Pred. No. 2.7e-136;
Matches 290; Conservative   0; Mismatches   0; Indels   0; G

Qy   462 GGACGGAGCGCAGGCAGGTGTGACGGACAGTGAGTGCTGGGAGGAGCCAGAJ
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9 GGACGGAGCGCAGGCAGGTGTGACGGACAGTGAGTGCTGGGAGGAGCCAGAJ

Qy   522 CAGCTCCAGCCCTCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGC
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   69 CAGCTCCAGCCCTCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGC

Qy   582 GCTCTACTACCTGTACTGTCAGGTGCACTTTGATGAGGGGAAGGCTGTCTACCTGJ
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   129 GCTCTACTACCTGTACTGTCAGGTGCACTTTGATGAGGGGAAGGCTGTCTACCTGJ

Qy   642 GGACTTGCTGGTGGATGGGTGCTGGCCCTGCGCCGTGCCTGGGAGGAATTCTCAGCCJ
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   189 GGACTTGCTGGTGGATGGGTGCTGGCCCTGCGCCGTGCCTGGGAGGAATTCTCAGCCJ

Qy   702 GGCCAGTTCCCTCGGGCCCCAGCTCCGCCTCTGCCAGGTGTCTGGGCTGT 751
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   249 GGCCAGTTCCCTCGGGCCCCAGCTCCGCCTCTGCCAGGTGTCTGGGCTGT 298

RESULT 48
BG686319
LOCUS   BG686319   587 bp   mRNA   linear   EST 0.

---

mer: M13 FORWARD
es.
          Location/Qualifiers
          1..333
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="UI-E-CL1-afe-m-20-0-UI"
                  /tissue_type="human retina"
                  /dev_stage="adult"
                  /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                  /clone_lib="UI-E-CL1"
                  /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                  modified polylinker; Site_1: EcoR I; Site_2: Not I;
                  UI-E-CL1 is a normalized cDNA library containing the
                  following tissue(s): retina. The library was constructed
                  according to Bonaldo, Lennon and Soares, Genome Research,
                  6:791-806, 1996. First strand cDNA synthesis was primed
                  with an oligo-dT primer containing a Not I site. Double
                  stranded cDNA was ligated to an EcoR I adaptor, digested
                  with Not I, and cloned directionally into pT7T3-Pac
                  vector. The oligonucleotide used to prime the synthesis of
                  first-strand cDNA contains a library tag sequence that is
                  located between the Not I site and the (dT)18 tail. The
                  sequence tag for this library is CCGCG. This library was
                  created for the program, Gene Discovery in the Visual
                  System, supported by National Eye Institute (NEI).
                  TAG_TISSUE=human retina
                  TAG_LIB=UI-E-CL1
                  TAG_SEQ=CCGCG"

arity   21.6%;   Score 296;   DB 13;   Length 333;
onservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

CCAATCCCTGACCCTTGAGGCCCCCAGTGAATCTCGACTCCCCCCTGCCACAGAC 1137
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCAATCCCTGACCCTTGAGGCCCCCAGTGAATCTCGACTCCCCCCTGCCACAGAC 274

AGGGCCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCCCACTT 1197
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGGGCCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCCCACTT 214

CACTAAGAGGGGCTGGACCTGGGCGGCAGGAAGCCAAAGAGACTGGGCCTAGGCCAG 1257
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CACTAAGAGGGGCTGGACCTGGGCGGCAGGAAGCCAAAGAGACTGGGCCTAGGCCAG 154

TCCCAAATGTGAGGGGCCAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTG 1317
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCCCAAATGTGAGGGGCCAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTG 94

TTTAAAACAGATATTATTTTTTATTATTATTGTGACAAAATGTTGATAAATGG 1373
|||||||||||||||||||||||||||||||||||||||||||||||||||||
TTTAAAACAGATATTATTTTTTATTATTATTGTGACAAAATGTTGATAAATGG 38

946   298 bp   mRNA   linear   EST 28-FEB-2002
CQ1-acs-a-05-0-UI.r1 UI-E-CQ1 Homo sapiens cDNA clone
CQ1-acs-a-05-0-UI 5', mRNA sequence.
946.
946.1 GI:19002204

sapiens (human)
sapiens
yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
lia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ases 1 to 298)
do,M.F., Lennon,G. and Soares,M.B.
lization and subtraction: two approaches to facilitate gene
very

```
Db   377 AAGTGCACCTAAAGGCCGGAAAACACCGGCTCGAAGAGCGATCGCAGCCATTATGA
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   444 TCATCCACGACCTGGACAGGACCGGACGGAGCCAGGCAG 478
             ||||||||||||||||||||||||||||||||||||||
Db   437 TCATCCACGACCTGGACAGGACCGGAGCCGAGCCAGGCAG 471
```

RESULT 49
AF163779
LOCUS   AF163779   1027 bp   DNA   linear   GSS 29-
DEFINITION AF163779 Human Homo sapiens genomic clone BAC750E14, genomi
            sequence.
ACCESSION   AF163779
VERSION   AF163779.1 GI:5726439
KEYWORDS   GSS.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelec
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1027)
  AUTHORS   Cousin,P., Billotte,J., Chaubert,P. and Shaw,P.H.
  TITLE   Physical map of 17p13 and the genes adjacent to p53
  JOURNAL   Genomics 63 (1), 60-68 (2000)
  MEDLINE   20130114
  PUBMED   10662545
COMMENT   Contact: Shaw PH
            Experimental Oncology
            Institute of Pathology
            Rue du Bugnon 25, Lausanne, VD 1011, Switzerland
            sub_clone=AB2R Asc-BamHI PSL1180
            Class: BAC subclone.
FEATURES   Location/Qualifiers
     source   1..1027
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /map="17p"
                     /clone="BAC750E14"
                     /clone_lib="Human"
ORIGIN

Query Match   19.7%; Score 271; DB 28; Length 1027;
Best Local Similarity 99.7%; Pred. No. 1.7e-126;
Matches 321; Conservative 0; Mismatches 1; Indels 0; G.

Qy 1052 CCACCGCCACTCTCCACCTCACTAGCTCCCAATCCTCAGCCCTTTGAGGCCCCC.
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   268 CCACCGCCACCTCTCCACCTCCACTCACTAGCTCCCAATCCTCAGCCCTTTGAGGSCCCC

Qy 1112 ATCTGACTCCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGT
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   328 ATCTGACTCCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGT

Qy 1172 AAGGATGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGC
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   388 AAGGATGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGC

Qy 1232 AGCCAAAGAGACTGGGCTCAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAA
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   448 AGCCAAAGAGACTGGGCTCAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAA

Qy 1292 AGCTCCTCCCTTGAGAATTCCTGGATTTTTAAAACAGATATTATTTTTTATTAT
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   508 AGCTCCTCCCTTGAGAATTCCTGTGGATTTTTAAAACAGATATTATTTTTATTAT

Qy 1352 GTGACAAAATGTTGATAAATGG 1373
             |||||||||||||||||||||
Db   568 GTGACAAAATGTTGATAAATGG 589

RESULT 50
AI760777/c
```

---

32F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4766071 5',
quence.
9
9.1  GI:13917716

piens (human)
piens
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 587)
  http://mgc.nci.nih.gov/.
l Institutes of Health, Mammalian Gene Collection (MGC)
shed (1999)
: Robert Strausberg, Ph.D.
  cgapbs-r@mail.nih.gov
Procurement: Louis M. Staudt, M.D., Ph.D.
ibrary Preparation: Ling Hong/Rubin Laboratory
ibrary Arrayed by: The I.M.A.G.E. Consortium (LLNL)
quencing by: Incyte Genomics, Inc.
distribution: MGC clone distribution information can be
hrough the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov
LLCM1625 row: p  column: 08
ality sequence stop: 587.
          Location/Qualifiers
     1..587
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4766071"
                     /tissue_type="primary B-cells from tonsils (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_48"
                     /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by Ling
                     Hong in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."

urity      20.6%;  Score 283;  DB 12;  Length 587;
99.6%;  Pred. No. 1.2e-132;
onservative  0;  Mismatches  1;  Indels  1;  Gaps  1;

CCGGCGGCTCCCCCCGATCCTCGGGTCCCGGGATGGGGGGCGCGGGGGCGGGGGGCCGGGTGAGG 84
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCGGCGGCTCCCCCCGATCCTCCGGGTCCCGGGATGGGGGGCGCGGGGGCGGGGGGCCGGGTGAGG 76

ACAGCCCCCGCGCCCCCCATGGCCGCCGCCGTCGGAGCCAGAGGCGGCAGAGGCGGAGGGGCGCCGG 144
  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ACAGCCCCCGCGCCCCCCATGGCCGCCGCCGTCGGAGCCAGAGGCGGCAGAGGCGGAGGGGCGCCGG 136

AGCCCGGGCACCGCCCCTGCTCGTGTCCCGCTCGCGCTGGGCTGGGCCTGGGC-GCTGGC 203
  |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
AGCCGGGCACCGCCCCTGCTCGTGTCCCGCTCGCGCTGGGCTGGGCCTGGGCTGCTGGC 196

CTCGGCCTCCTGCTGGCCGTGTGTCAGTTGGGGAGCCGGGCATCGCTGTCTCCGCCCA 263
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTCGGCCTCCTGCTGGCCGTCAGTTGGGGAGCCGGGCATCGCTGTCTCCGCCCA 256

CCTGCCCAGGAGGAGTGGTGGCAGAGGAGGACCAGGACCCGTCGGAACTGAATCC 323
  |||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCTGCCCAGGAGGAGTGGTGGCAGAGGAGGACCAGGACCCGTCGGAACTGAATGC 316

ACAGAAGAAAGCCAGGATCCTCGCCTTTCCTGAACCGACTAGTTCGGCCTTCGCAG 383
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ACAGAAGAAAGCCAGGATCCTGCGCCTTTCCTGAACCGACTAGTTCGGCCTTCGCGCAG 376

GCACCTAAAGGCCGGAAAACACGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGT 443

77
5.xl NCI_CGAP_Kid12 Homo sapiens cDNA clone IMAGE:2398377 3', mRNA    linear    EST 24-JUN-1999    346 bp
sequence.
77
77.1    GI:5176444
apiens (human)
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 346)
AP http://www.ncbi.nlm.nih.gov/ncicgap
al Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Gene Index
ished (1997)
t: Robert Strausberg, Ph.D.
cgapbs-r@mail.nih.gov
Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
-Buck, M.D., Ph.D.
Library Preparation: M. Bento Soares, Ph.D
Library Arrayed by: Greg Lennon, Ph.D
equencing by: Washington University Genome Sequencing Center
distribution: NCI-CGAP clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL at:
o.llnl.gov/bbrp/image/image.html
imer: -40UP from Gibco.
        1..346
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:2398377"
        /tissue_type="2 pooled tumors (clear cell type)"
        /lab_host="DH10B"
        /clone_lib="NCI_CGAP_Kid12"
        /note="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with
        a modified polylinker; Site 1: Not I; Site 2: Eco RI;
        Plasmid DNA from the normalized library NCI_CGAP_Kid5 was
        prepared, and ss circles were made in vitro. Following HAP
        purification, this DNA was used as tracer in a subtractive
        hybridization reaction. The driver was PCR-amplified cDNAs
        from a pool of 5,000 clones made from the same library
        (cloneIDs 1323912-1325831, 1471368-1472903 and
        1492104-1493255). Subtraction by Bento Soares and M.
        Fatima Bonaldo. "

larity      19.4%;    Score 267;    DB 9;    Length 346;
       99.7%;    Pred. No. 1.5e-124;
onservative  0;    Mismatches  1;    Indels  0;    Gaps  0;

CCACTCTCCACCTCACTAGCTCCCCAATCCCTGACCCTTGAGGCCCCCAGTGATCT 1115
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCACTCTCCACCTCACTAGCTCCCCAATCCCTGACCCTTGAGGCCCCCAGTGATCT 287

CTCCCCCCTGGCCACAGACCCCCAGGGCATTGTTCACTGTACTCTGTGGGCAAGG 1175
|||||||||||||||||||| |||||||||||||||||||||||||||||||
CTCCCCCCTGGCCACAGACCCCNCAGGGCATTGTTCACTGTACTCTGTGGGCAAGG 227

GGTCCAGAAGAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGCGGCAGGAAGCC 1235
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGTCCAGAAGAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGCGGCAGGAAGCC 167

GAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGGAGAAACAAGACAAGCT 1295
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGGAGAAACAAGACAAGCT 107

CCCTTGAGAATTCCCTGTGGATTTTAAAACAGATATTTTTATTATTATTGTGA 1355
||||||||||||||||||||||| ||||||||||||||||||||||||||
CCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTTTTATTATTATTGTGA 47

AATGTTGATAAATGG 1373
|||||||||||||||

---

Db      46  CAAAATGTTGATAAATGG 29

RESULT 51
BF940141/c
LOCUS        BF940141        346 bp    mRNA    linear    EST 22
DEFINITION   nac68g06.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:3
             similar to contains element MSR1 repetitive element ;, mRN
             sequence.
ACCESSION    BF940141
VERSION      BF940141.1    GI:12357461
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 346)
AUTHORS      NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap
TITLE        National Cancer Institute / National Institute of Neurolog
             Disorders and Stroke, Brain Tumor Genome Anatomy Project
             (CGAP/BTGAP), Tumor Gene Index
JOURNAL      Unpublished (1998)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfe
             Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fati
             Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencir
             Clone distribution: NCI-CGAP clone distribution informati
             found through the I.M.A.G.E. Consortium/LLNL, send email t
             info@image.llnl.gov
             Seq primer: -40UP from Gibco
             High quality sequence stop: 321.
FEATURES     Location/Qualifiers
     source  1..346
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:3439667"
             /tissue_type="glioblastoma (pooled)"
             /lab_host="DH10B"
             /clone_lib="NCI_CGAP_Brn23"
             /note="Organ: brain; Vector: pT7T3D-Pac (Pharmac
             modified polylinker; Site 1: Not I; Site 2: Eco
             strand cDNA was primed with a Not I - oligo(dT)
             TGTTACCAATCTGAAGTGGGAGCGGCGCGCATATCTTTTTTTTTTTTT
             T 3'|; double-stranded cDNA was ligated to Eco
             adaptors (Pharmacia), digested with Not I and c
             the Not I and Eco RI sites of the modified pT7T
             Library is normalized, and was constructed by B
             Soares and M.Fatima Bonaldo."

ORIGIN

Query Match          19.1%;    Score 262;    DB 10;    Length 346;
Best Local Similarity    100.0%;    Pred. No. 5.4e-122;
Matches 262;    Conservative   0;    Mismatches   0;    Indels   0;

Qy  1048  TCCCCCACCGCCACTCTCCACCTCACTAGCTCCCCAATCCCTGACCCTTGAGG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   346  TCCCCCACCGCCACTCTCCACCTCACTAGCTCCCCAATCCCTGACCCTTGAGG

Qy  1108  AGTGATCTCGACTCCCCCCTGGCCACAGACCCCCAGGGCATTGTTCACTGTAC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   286  AGTGATCTCGACTCCCCCCTGGCCACAGACCCCCAGGGCATTGTTCACTGTAC

Qy  1168  GGGCAAGGATGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTG
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   226  GGGCAAGGATGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTG

Qy  1228  AGGAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCAGAGTTCCCAAATGTGAGGGGCGAGA
          ||||||||||||||||||||||||||||||||||||||||||||||||||||

```
GCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCCAGAAACAA 107

GCTCCTCCCTTGAGAAT 1309
|||||||||||||||||
GCTCCTCCCTTGAGAAT 85
```

```
8
2 .x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:3441742 3'
to TR:Q9UK76_Q9UK76 HN1 PROTEIN .; mRNA sequence.
8
8.1  GI:12512043
```

```
piens (human)
piens
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 561)
DS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
l Cancer Institute / National Institute of Neurological
rs and Stroke, Brain Tumor Genome Anatomy Project
TGAP), Tumor Gene Index
shed (1998)
: Robert Strausberg, Ph.D.
cgapbs-r@mail.nih.gov
Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
```

```
ibrary Preparation: M. Bento Soares, Ph.D., M. Fatima
, Ph.D.
ibrary Arrayed by: Greg Lennon, Ph.D.
quencing by: Washington University Genome Sequencing Center
distribution: NCI-CGAP clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL, send email to:
age.llnl.gov
mer: -40UP from Gibco
ality sequence stop: 260.
        Location/Qualifiers
        1..561
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:3441742"
        /tissue_type="glioblastoma (pooled)"
        /lab_host="DH10B"
        /clone_lib="NCI_CGAP_Brn23"
        /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
        modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
        strand cDNA was primed with a Not I - oligo(dT) primer (5'
        TGTTACCAATCTGAAGTGGGAGCGGCCGCCATAtCTTTTTTTTTTTTTTTTTTTT
        T 3'); double-stranded cDNA was ligated to Eco RI
        adaptors (Pharmacia), digested with Not I and cloned into
        the Not I and Eco RI sites of the modified pT7T3 vector.
        Library is normalized, and was constructed by Bento
        Soares and M.Fatima Bonaldo."
```

```
arity   18.0%;   Score 247;  DB 10; Length 561;
arity  100.0%;  Pred. No. 2.7e-114;
onservative  0; Mismatches  0; Indels  0; Gaps  0;

CACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGA 1186
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGA 212

CCCACTTCAGGCACTAAGAGGGGGCTGGACCTGGCGGCAGGAGAGCCAAAGAGACTGG 1246
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCACTTCAGGCACTAAGAGGGGGCTGGACCTGGCGGCAGGAGAGCCAAAGAGACTGG 152

AGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAG 1306
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAG 92
```

Right column:

```
1307 AATTCCCTGTGGATTTTTAAAACAGATATTATTTTTATTATTATTGTGACAAAATGT
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  91 AATTCCCTGTGGATTTTTAAAACAGATATTATTTTTATTATTATTGTGACAAAATGT

1367 TAAATGG 1373
     |||||||
  31 TAAATGG 25
```

```
RESULT 53
AW081731/c
LOCUS       AW081731              318 bp    mRNA    linear    EST 14-
DEFINITION  xb70a02.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:2581610 3' similar to contains element MSR1 repetiti
            element ; mRNA sequence.
ACCESSION   AW081731
VERSION     AW081731.1    GI:6036883
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 318)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project ((
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
  COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contacl
            IMAGE Consortium (info@image.llnl.gov) for further informal
            Seq primer: -40UP from Gibco
            High quality sequence stop: 314.
FEATURES    Location/Qualifiers
  source    1..318
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:2581610"
            /lab_host="DH10B"
            /clone_lib="Soares_NFL_T_GBC_S1"
            /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmac
            a modified polylinker; Site_1: Not I; Site_2: Eco
            Equal amounts of plasmid DNA from three normalize
            libraries (fetal lung NbHL19W, testis NHT, and B-
            NCI_CGAP_GCB1) were mixed, and ss circles were ma
            vitro. Following HAP purification, this DNA was u
            tracer in a subtractive hybridization reaction. T
            was PCR-amplified cDNAs from pools of 5,000 clone
            from the same 3 libraries. The pools consisted of
            I.M.A.G.E. clones 297480-302087, 682632-687239,
            726408-728711, and 729096-731399. Subtraction by
            Soares and M. Fatima Bonaldo."
```

```
ORIGIN
Query Match      17.2%;    Score 236;   DB 9; Length 318;
Best Local Similarity  100.0%;  Pred. No. 9.9e-109;
Matches  236;  Conservative   0;  Mismatches   0;  Indels   0;  G

Qy  1138 CCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGTCCAGAAGACCCC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   254 CCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGTCCAGAAGACCCC

Qy  1198 CAGGCACTAAGAGGGGGCTGGACCTGGCGGCAGGAGAAGCCAAAGAGACTGGGCCTAGG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   194 CAGGCACTAAGAGGGGGCTGGACCTGGCGGCAGGAGAAGCCAAAGAGACTGGGCCTAGG

Qy  1258 GAGTTCCCAAATGTGAGGGGCGAGAAACAAGCAAGCTCCTCCCTTGAGAATTCCC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   134 GAGTTCCCAAATGTGAGGGGCGAGAAACAAGCAAGCTCCTCCCTTGAGAATTCCC

Qy  1318 GATTTTTAAAACAGATATTATTTTATTATTATTGTGACAAAATGTGATAAATGC
```

```
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TTTAAAACAGATATTATTTTTATTATTATTGTGACAAAATGTTGATAAATGG 19
```

DEFINITION   wj17g05.x1 NCI_CGAP_Kid12 Homo sapiens cDNA clone IMAGE:24
             similar to contains element PTR5 repetitive element ;. mRN
             sequence.
ACCESSION   AI863563
VERSION   AI863563.1 GI:5527670
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 253)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (
             Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Mic
             Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencin
             Clone distribution: NCI-CGAP clone distribution informati
             found through the I.M.A.G.E. Consortium/LLNL at:
             www-bio.llnl.gov/bbrp/image/image.html
             Insert Length: 669   Std Error: 0.00
             Seq primer: -40UP from Gibco
             High quality sequence stop: 253.
FEATURES   Location/Qualifiers
     source   1..253
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2403128"
                     /tissue_type="2 pooled tumors (clear cell type)"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Kid12"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharma
                     a modified polylinker; Site_1: Not I; Site_2: Ec
                     plasmid DNA from the normalized library NCI_CGAP
                     prepared, and ss circles were made in vitro. Fol
                     purification, this DNA was used as tracer in a s
                     hybridization reaction. The driver was PCR-ampli
                     from a pool of 5,000 clones made from the same 1
                     (cloneIDs 1323912-1325831, 1471368-1472903 and
                     1492104-1493255). Subtraction by Bento Soares an
                     Fatima Bonaldo. "

ORIGIN

Query Match   16.7%; Score 229; DB 9; Length 253;
Best Local Similarity   100.0%; Pred. No. 3.5e-105;
Matches 229; Conservative   0; Mismatches   0; Indels   0;

Qy   1145 GCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCCCACTTCA
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   253 GCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCCCACTTCA

Qy   1205 TAAGAGGGCTGGACCTGGCGCGGCAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGA
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   193 TAAGAGGGCTGGACCTGGCGCGGCAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGA

Qy   1265 CAAATGTGAGGGGCCAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGA
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   133 CAAATGTGAGGGGCCAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGA

Qy   1325 AAAACAGATATTATTTTTATTATTATTGTGACAAAATGTTGATAAATGG 1373
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   73 AAAACAGATATTATTTTTATTATTATTGTGACAAAATGTTGATAAATGG 25

RESULT 56
AI682487/c
LOCUS   AI682487   238 bp   mRNA   linear   EST 1
```

```
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TTTAAAACAGATATTATTTTTATTATTATTGTGACAAAATGTTGATAAATGG 19
```

             264 bp   mRNA   linear   EST 06-APR-2000

41   0.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
2978587 3', mRNA sequence.
41
41.1 GI:7454367

apiens (human)
sapiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 264)
AP http://www.ncbi.nlm.nih.gov/ncicgap.
al Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Gene Index
ished (1997)
t: Robert Strausberg, Ph.D.
cgapbs-r@mail.nih.gov
clone is available royalty-free through LLNL ; contact the
Consortium (info@image.llnl.gov) for further information.
imer: -40UP from Gibco
quality sequence stop: 263.
Location/Qualifiers
1..264
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:2978587"
        /lab_host="DH10B"
        /clone_lib="Soares_NFL_T_GBC_S1"
        /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
        a modified polylinker; Site_1: Not I; Site_2: Eco RI;
        Equal amounts of plasmid DNA from three normalized
        libraries (fetal lung NbHL19W, testis NHT, and B-cell
        NCI_CGAP_GCBl) were mixed, and ss circles were made in
        vitro. Following HAP purification, this DNA was used as
        tracer in a subtractive hybridization reaction. The driver
        was PCR-amplified cDNAs from pools of 5,000 clones made
        from the same 3 libraries. The pools consisted of
        I.M.A.G.E. clones 297480-302087, 682632-687239,
        726408-728711, and 729096-731399. Subtraction by Bento
        Soares and M. Fatima Bonaldo. "

larity   16.8%;   Score 230; DB 10; Length 264;
        100.0%;   Pred. No. 1.1e-105;
Conservative   0; Mismatches   0; Indels   0; Gaps   0;

ATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCCCACTTCAGGCA 1203
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCCCACTTCAGGCA 195

AGAGGGGCTGGACCTGGCGCGGCAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTC 1263
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGAGGGGCTGGACCTGGCGCGGCAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTC 135

AATGTGAGGGGCCAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTT 1323
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AATGTGAGGGGCCAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTT 75

AACAGATATTATTTTTATTATTATTGTGACAAAATGTTGATAAATGG 1373
|||||||||||||||||||||||||||||||||||||||||||||||
AACAGATATTATTTTTATTATTATTGTGACAAAATGTTGATAAATGG 25

563   253 bp   mRNA   linear   EST 21-DEC-1999
```

.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:2324405 3',
quence.
7
7.1 GI:4892669
piens (human)
piens
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 238)
P http://www.ncbi.nlm.nih.gov/ncicgap.
l Cancer Institute, Cancer Genome Anatomy Project (CGAP),
ene Index
shed (1997)
: Robert Strausberg, Ph.D.
cgapbs-r@mail.nih.gov
Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
Buck, M.D., Ph.D.
ibrary Preparation: M. Bento Soares, Ph.D.
ibrary Arrayed by: Greg Lennon, Ph.D.
quencing by: Washington University Genome Sequencing Center
distribution: NCI-CGAP clone distribution information can be
hrough the I.M.A.G.E. Consortium/LLNL at:
.llnl.gov/bbrp/image/image.html
Length: 223   Std Error: 0.00
mer: -40UP from Gibco.
Location/Qualifiers
1..238
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2324405"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Pr28"
/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
with a modified polylinker; Plasmid DNA from the
normalized library NCI_CGAP_Pr22 was prepared, and ss
circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneIDs
985608-986759, 1101192-1101959, and 1217928-1220615).
Subtraction by Bento Soares and M. Fatima Bonaldo."

rity   16.5%;   Score 227;   DB 9;   Length 238;
100.0%;   Pred. No. 3.6e-104;   Indels   0;   Gaps   0;
onservative   0;   Mismatches   0;

IGTTCACTGTACTCTGTGGGCAAGGATGGTCAGAGAGACCCCACTTCAGGCACTA 1206
IGTTCACTGTACTCTGTGGGCAAGGATGGTCAGAGAGACCCCACTTCAGGCACTA 179

3GGCTGGACCTGGCGGCAGGAAGCCAAAGAGACTGGGCTAGGCCAGGAGTTCCCA 1266
3GGCTGGACCTGGCGGCAGGAAGCCAAAGAGACTGGGCTAGGCCAGGAGTTCCCA 119

TGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAA 1326
TGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAA 59

GATATATTTTTATTATTATTGTGACAAAATGTTGATAAATCG 1373
GATATATTTTTATTATTATTGTGACAAAATGTTGATAAATCG 12

50  633F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5213480 5',
equence.

---

ACCESSION   BI906850
VERSION     BI906850.1  GI:16169619
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 698)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (M
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11536 row: b column: 09
            High quality sequence stop: 613.
FEATURES    Location/Qualifiers
     source 1..698
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5213480"
            /tissue_type="leukocyte"
            /clone_lib="NIH_MGC_118"
            /lab_host="DH10B"
            /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2:
            (destroyed); RNA source leukocytes from anonymous
            non-activated adult donors. Library is oligo-dT
            and directionally cloned (EcoRV site is destroye
            cloning). Average insert size 1.7 kb, insert siz
            1.2-3.3 kb. Library is normalized and enriched f
            full-length clones and was constructed by C. Grub
            (invitrogen). Research Genetics tracking code 02
            this is a NIH_MGC Library."
ORIGIN

Query Match   15.8%;   Score 217;   DB 12;   Length 698;
Best Local Similarity   100.0%;   Pred. No. 5.8e-99;
Matches 217;   Conservative   0;   Mismatches   0;   Indels   0;   G

Qy 262 CAGGAGCCTGCCCAGGAGGAGCTGGTGGCAGGAGGAGGACCAGGACCCGTCGGAACT
Db 235 CAGGAGCCTGCCCAGGAGGAGCTGGTGGCAGGAGGAGGACCAGGACCCGTCGGAACT

Qy 322 CCCCAGACAGAGAAAGCCAGGATCCTGCGGCCTTTCGTGACCGACTAGTTCGGCC
Db 295 CCCCAGACAGAGAAAGCCAGGATCCTGCGGCCTTTCGTGACCGACTAGTTCGGCC

Qy 382 AGAAGTGCACCTAAGGCCGAAAACACGGGCTCGAAGAGGATCGCAGCCCATTA
Db 355 AGAAGTGCACCTAAGGCCGAAAACACGGGCTCGAAGAGGATCGCAGCCCATTA

Qy 442 GTTCATCCACGACCTGGACAGGAGGACGAGGGCGCAGGCAG 478
Db 415 GTTCATCCACGACCTGGACAGGAGGACGAGGGCGCAGGCAG 451

RESULT 58
BG054914/c
LOCUS       BG054914   351 bp   mRNA   linear   EST 25
DEFINITION  nac92a11.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:3
            similar to contains element MSR1 repetitive element ;, mRN
            sequence.
ACCESSION   BG054914
VERSION     BG054914.1  GI:12512115
KEYWORDS    EST.
SOURCE      Homo sapiens (human)

VERSION     AI669243.1  GI:4834017
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 397)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (
            CGAP), Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Mi
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencin
            Clone distribution: NCI-CGAP clone distribution informati
            found through the I.M.A.G.E. Consortium/LLNL at:
            www.bio.llnl.gov/bbrp/image/image.html
            Insert Length: 438    Std Error: 0.00
            Seq primer: -40UP from Gibco.

FEATURES             Location/Qualifiers
     source          1..397
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2315088"
                     /sex="male"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Pr28"
                     /note="Organ: prostate; Vector: pT7T3D-Pac (Pharm
                     with a modified polylinker; Plasmid DNA from the
                     normalized library NCI_CGAP_Pr22 was prepared, an
                     circles were made in vitro. Following HAP purific
                     this DNA was used as tracer in a subtractive hybr
                     reaction. The driver was PCR-amplified cDNAs from
                     of 5,000 clones made from the same library (clone
                     985608-986759, 1101192-1101959, and 1217928-12206
                     Subtraction by Bento Soares and M. Fatima Bonald

ORIGIN

Query Match            15.1%;    Score 207;   DB 9;   Length 397;
Best Local Similarity  99.4%;    Pred. No. 6.6e-94;
Matches  307;   Conservative  0;   Mismatches  2;   Indels   0;

Qy   1065  TCCACCTCACTAGCTCCCCAATCCCTGACCCTTGAGGCCCCAGTGATCTGAC
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    343  TCCACCTCACTAGCTCCCCAATCCCTGACCCTTGAGGCCCCAGTGATCTGAC

Qy   1125  CCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGT
           ||||||||||||||||||||||||||||||||||||||||||
Db    283  CCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGT

Qy   1185  GAAGACCCCACTTCAGGCACTAAGAGGGGCTGACCTGGCGGCAGGAAGCCAAAG
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    223  GAAGACCCCACTTCAGGCACTAAGAGGGGCTGACCTGGCGGCAGGAAGCCAAAA

Qy   1245  GGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    163  GGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTC

Qy   1305  AGAATTCCCTGTGGATTTTTAAAACAGATATTTTATTATTATTGTGACAAA
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    103  AGAATTCCCTGTGGATTTTTAAAACAGATATTTTATTATTATTGTGACAAA

Qy   1365  GATAAATGG 1373
           |||||||||
Db     43  GATAAATGG 35

---

apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 351)
NDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
al Cancer Institute / National Institute of Neurological
ers and Stroke, Brain Tumor Genome Anatomy Project
BTGAP), Tumor Gene Index
ished (1998)
t: Robert Strausberg, Ph.D.
cgapbs-r@mail.nih.gov
Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Library Preparation: M. Bento Soares, Ph.D., M. Fatima
o, Ph.D.
Library Arrayed by: Greg Lennon, Ph.D.
equencing by: Washington University Genome Sequencing Center
distribution: NCI-CGAP clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL, send email to:
mage.llnl.gov
imer: -40UP from Gibco.
uality sequence stop: 339.

1..351
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3441693"
/tissue_type="glioblastoma (pooled)"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Brn23"
/note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCGAAGTGGGAGGCGGCCGCATATCTTTTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldo."

arity        15.4%;   Score 211;   DB 10;   Length 351;
onservative  99.6%;   Pred. No. 5.8e-96;
conservative  0;   Mismatches  0;   Indels  1;   Gaps   0;

CCCACCGCGCCACTCTCCACCTCCACTAGCTCCACTAGCTCCCCAATCCCTGACCCTTGAGGCCCCC  1107
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCACCGCGCCACTCTCCACCTCCACTAGCTCCACTAGCTCCCCAATCCCTGACCCTTGAGGCCCCC  292

GATCTCGACTCCCCCCTGGCCACAGACCCCCAGGGCCATTGTGTTCACTGTACTCTGT  1167
|||||||||||||||||||||||||||||||||||||||||||||||||||
GATCTCNGACTCCCCCCTGGCCACAGACCCCCAGGGCCATTGTGTTCACTGTACTCTGT  232

CAAGGATGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGC  1227
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAGGATGGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGC  172

AAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAA  1287
||||||||||||||||||||||||||||||||||||||||||||||||||||||
AAGCCAAAGAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAA  112

AAGCTCCTCCCTTGAGAAT  1309
|||||||||||||||||||
AAGCTCCTCCCTTGAGAAT  90

243   397 bp    mRNA    linear    EST 17-DEC-1999
01.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:2315088 3',
sequence.
243

4        894 bp    mRNA      linear   EST 16-OCT-2001
26F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5217367 5',
quence.
4
4.1  GI:16171193

piens (human)
piens
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 894):
l Institutes of Health, Mammalian Gene Collection (MGC)
shed (1999)
: Robert Strausberg, Ph.D.
cgapbs-r@mail.nih.gov
Procurement: Life Technologies, Inc.
ibrary Preparation: Life Technologies, Inc.
http://mgc.nci.nih.gov/,
quencing by: Incyte Genomics, Inc.
distribution: MGC clone distribution information can be
hrough the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov
LIAM11546 row: d  column: 08
ality sequence start: 5
ality sequence stop: 460.
         Location/Qualifiers
         1..894
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /clone="IMAGE:5217367"
         /tissue_type="leukocyte"
         /lab_host="DH10B"
         /clone_lib="NIH_MGC_118"
         /note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV
         (destroyed); RNA source leukocytes from anonymous pool of
         non-activated adult donors. Library is oligo-dT primed
         and directionally cloned  (EcoRV site is destroyed upon
         cloning).  Average insert size 1.7 kb, insert size range
         1.2-3.3 kb. Library is normalized and enriched for
         full-length clones and was constructed by C. Gruber
         (Invitrogen).  Research Genetics tracking code 027. Note:
         this is a NIH_MGC Library."

         15.0%;   Score 206;  DB 12;  Length 894;
urity   100.0%;  Pred. No. 2.5e-93;
nservative     0;  Mismatches    0;  Indels   0;  Gaps    0;

AGGACCAGGACCCGTCGGAACTGAATCCCCAGACAGAGAAAGCCAGGATCCTGCG 351
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGGACCAGGACCCGTCGGAACTGAATCCCCAGACAGAGAAAGCCAGGATCCTGCG 281

CCTGAACCGACTAGTTCGGCCTCGCAGAGTGCACCTAAAGGCCGGAAAACACGG 411
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCTGAACCGACTAGTTCGGCCTCGCAGAGTGCACCTAAAGGCCGGAAAACACGG 341

GAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACGGAGCG 471
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
GAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACGGAGCG 401

CAGGTGTGGACGGACAGTGAG 497
|||||||||||||||||||||
CAGGTGTGGACGGACAGTGAG 427

41        465 bp    mRNA      linear   EST 18-AUG-1998
5.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone


IMAGE:1651448 3' similar to contains MSR1.t3 MSR1 repetitiv
element ;, mRNA sequence.

ACCESSION   AI091441
VERSION     AI091441.1  GI:3430500
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelec
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 465)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project ((
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact
            IMAGE Consortium (info@image.llnl.gov) for further informat
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 465.
FEATURES             Location/Qualifiers
     source          1..465
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1651448"
                     /lab_host="DH10B"
                     /clone_lib="Soares_NSF_F8_9W_OT_PA_P_S1"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmaci
                     a modified polylinker; Site 1: Not I; Site 2: Eco
                     Equal amounts of plasmid DNA from five normalized
                     libraries were mixed, and ss circles were made in
                     Following HAP purification, this DNA was used as t
                     a subtractive hybridization reaction. The driver
                     PCR-amplified cDNAs from pools of 5,000 clones mat
                     the same 5 libraries. The pools consisted of the 1
                     libraries and cloneIds: Soares NbHSF pool 1:
                     309384-310919, 323208-325895 Soares Nb2HP pool 1:
                     145032-147335, 147720-148103, 148872-149255, 1500:
                     150407, 151176-152327 Soares Nb2HF8-9W pool 1:
                     758280-760583, 772104-774407 Soares NbHPA pool 1:
                     304776-306311, 320136-322823, 326280-326663 Soare:
                     pool 1: 723720-726407, 739080-740999 Subtraction
                     Soares and M. Fatima Bonaldo."

ORIGIN

     Query Match      14.7%;  Score 202;  DB 9;  Length 465;
     Best Local Similarity 100.0%; Pred. No. 2.4e-91;
     Matches 202; Conservative   0; Mismatches    0;  Indels    0;  G

Qy     25 TCGGCCGCGGCGCTCCCCCTCCCCCCGATCCTCGGGTCCCGGATGGGGGGCGGT
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    183 TCGGCCGCGGCGCTCCCCCTCCCCCCGATCCTCGGGTCCCGGATGGGGGGCGGT

Qy     85 CAGGCACAGCCCCCGCCCCGCCCCCATGGCGCCCCGTCGGAGGCGCAGAGGCGGAGGGGCG
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    243 CAGGCACAGCCCCCGCCCCGCCCCCATGGCGCCCCGTCGGAGGCGCAGAGGCGGAGGGGCG

Qy    145 GGGGAGCCGGGGACCGCCTGCTCGTCCGCTCGGCTGGGCCTGGGCCTGGGCCTGGGCCTT
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    303 GGGGAGCCGGGGACCGCCTGCTCGTCCGCTCGGCTGGGCCTGGGCCTGGGCCTGGGCGCT

Qy    205 TGCCTCGGCCTCCTGCTGGCCG 226
          ||||||||||||||||||||||
Db    363 TGCCTCGGCCTCCTGCTGGCCG 384

RESULT 62
BF222608/c
LOCUS       BF222608         224 bp    mRNA      linear   EST 09
DEFINITION  7p56d12.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:364
            mRNA sequence.
ACCESSION   BF222608

08.1  GI:11129886
apiens (human)
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 224)
AP http://www.ncbi.nlm.nih.gov/ncicgap.
al Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Gene Index
ished (1997)
t: Robert Strausberg, Ph.D.
   cgapbs-r@mail.nih.gov
Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
-Buck, M.D. Ph.D.
Library Preparation: M. Bento Soares, Ph.D.
Library Arrayed by: Greg Lennon, Ph.D.
sequencing by: Washington University Genome Sequencing Center
distribution: NCI-CGAP clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL, send email to:
mage.llnl.gov.
      Location/Qualifiers
      1..224
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:3649966"
        /sex="male"
        /dev_stage="adult"
        /lab_host="DH10B"
        /clone_lib="NCI_CGAP_Pr28"
        /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
        with a modified polylinker; Plasmid DNA from the
        normalized library NCI_CGAP_Pr22 was prepared, and ss
        circles were made in vitro. Following HAP purification,
        this DNA was used as tracer in a subtractive hybridization
        reaction. The driver was PCR-amplified cDNAs from a pool
        of 5,000 clones made from the same library (cloneIDs
        985608-986759, 1101192-1101959, and 1217928-1220615).
        Subtraction by Bento Soares and M. Fatima Bonaldo."

larity    13.9%;   Score 191;   DB 10;   Length 224;
Conservative  100.0%;  Pred. No. 8.5e-86;
Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

AAGACCCCACTTCAGGCACTAAGAGAGGGCTGGACCTGGCGGCCAGGAAGCCAAGAGA 1242
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AAGACCCCACTTCAGGCACTAAGAGAGGGCTGGACCTGGCGGCCAGGAAGCCAAGAGA 165

GGCCTAGGCCAGGAAGTTCCCAAATGTGAGGGGGCGAGAAACAAGACAAGCTCCTCCCT 1302
||||||||||||||||||||||||||||||||||||||||||||||||||
GGCCTAGGCCAGGAAGTTCCCAAATGTGAGGGGGCGAGAAACAAGACAAGCTCCTCCCT 105

GAATTCCCTGTGGATTTTTAAAACAGATATTATTTTTATTATTATTGTGACAAAATG 1362
||||||||||||||||||||||||||||||||||||||||||||||||
GAATTCCCTGTGGATTTTTAAAACAGATATTATTTTTATTATTATTGTGACAAAATG 45

ATAAATGG 1373
||||||||
ATAAATGG 34

4    282 bp   mRNA   linear   EST 31-DEC-2002
.0003761 Human adult (K.Okubo) Homo sapiens cDNA, mRNA
ence.
4.1  GI:1433224
sapiens (human)
sapiens

---

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 282)
REFERENCE
AUTHORS   Okubo,K.
TITLE     BodyMap; human gene expression database
JOURNAL   Unpublished (1995)
COMMENT   Contact: Okubo,K.
          Institute for Molecular and Cellular Biol
          Osaka University
          1-3,Yamada-oka, Suita, Osaka Pref. 565, Japan
          Tel: 06-877-5111(ex.3315)
          Email: kousaku@imcb.osaka-u.ac.jp
          We are not submitting the same cDNA sequence redundantly t
          since 1993. For the abundance information of clones with t
          sequence in this library and as well as in other 3'-direct
          libraries, see ' http://www.imcb.osaka-u.ac.jp/bodymap'.
          sequences of the clones represented by this GS sequences
          found there.
FEATURES  Location/Qualifiers
source    1..282
          /organism="Homo sapiens"
          /mol_type="mRNA"
          /db_xref="taxon:9606"
          /dev_stage="adult"
          /clone_lib="Human adult (K.Okubo)"
          /note="One or more human adult tissue"
ORIGIN

Query Match    13.3%;   Score 182;   DB 13;   Length 282;
Best Local Similarity  100.0%;   Pred. No. 3.5e-81;
Matches 182;   Conservative   0;   Mismatches   0;   Indels   0;   (

Qy  1161 ACTCTGTGGCAAGGATGGGTCAGAAGACCCCACTTCAGGCACTAAGAGGGGCT
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    53 ACTCTGTGGCAAGGATGGGTCAGAAGACCCCACTTCAGGCACTAAGAGGGGCT

Qy  1221 TGGCGGCAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGG
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   113 TGGCGGCAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGG

Qy  1281 GAAACAAGACAAGCTCCTCCCTTGAGAATTCCCGTGGATTTTTAAAACAGATAT
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   173 GAAACAAGACAAGCTCCTCCCTTGAGAATTCCCGTGGATTTTTAAAACAGATAT

Qy  1341 TT 1342
         ||
Db   233 TT 234

RESULT 64
CF994566
LOCUS      CF994566           1064 bp    mRNA    linear    EST 2
DEFINITION AGENCOURT 15621154 NIH_MGC_147 Homo sapiens cDNA clone
           IMAGE:30520331 5', mRNA sequence.
ACCESSION  CF994566
VERSION    CF994566.1  GI:38510626
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutel
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
REFERENCE  1 (bases 1 to 1064)
AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection
JOURNAL    Unpublished (1999)
COMMENT    Contact: Daniela S. Gerhard, Ph.D.
           Office of Cancer Genomics
           National Cancer Institute / NIH
           Bldg. 31 Rm10A07 Bethesda, MD 20892
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: Dr. Stefan Hansson
           cDNA Library Preparation: Michael J. Brownstein (NHGRI)
           and advice from Piero Carninci (RIKEN)

## Left section

ibrary Arrayed by: The I.M.A.G.E. Consortium (LLNL)
quencing by: Agencourt Bioscience Corporation
distribution: MGC clone distribution information can be
hrough the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov
NDAM593 row: e   column: 12
ality sequence stop: 249.
Location/Qualifiers
1..1064
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:30520331"
        /tissue_type="Human Placenta"
        /lab_host="DH10B TonA"
        /clone_lib="NIH MGC 147"
        /note="Organ: placenta; Vector: pBluescriptR; Site 1:
        all-XhoI; Site 2: BamH; Oligo-dT primed using primer
        5'-TTTTTTTTTTTTTTVN-3', size-selected for average
        insert size 2.3 kb and normalized to ROT 5. This is a
        primary library enriched for full-length clones and
        constructed using the Cap-trapper method (Carninci, in
        preparation). Library constructed by M. Brownstein
        (NIMH/NHGRI, National Institutes of Health). Note: This is
        a NIH_MGC library."

rity      12.7%;   Score 175;   DB 14;   Length 1064;
          100.0%;  Pred. No. 1.7e-77;
nservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

CCCCGCGCCCCCATGGCCGCCCGTCGGAGCCAGAGCGCGGAGGGGGCGCCGGGGGA 149
CCCCGCGCCCCCATGGCCGCCCGTCGGAGCCAGAGCGCGGAGGGGGCGCCGGGGGA 87
GCACCGCCGCCTGCTCGTCGTCCGCGCTCGGGCCTGGGCCTGGGCCTGGCTGCCT 209
GCACCGCCGCCTGCTCGTCGTCCGCGCTCGGGCCTGGGCCTGGGCCTGGCTGCCT 147
TCCTGCTGGCCGTGTCAGTTTGGGAGGCGGGCATCGTGTCCGCCAG 264
TCCTGCTGGCCGTGTCAGTTTGGGAGGCGGGCATCGTGTCCGCCAG 202

RT 15623743 NIH MGC 147 Homo sapiens cDNA clone
               1319 bp    mRNA    linear    EST 26-SEP-2003
RT 15623743 NIH_MGC_147 Homo sapiens cDNA clone
0527869 5', mRNA sequence.
33.1   GI:36348525
apiens (human)
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 1319)
: http://mgc.nci.nih.gov/.
al Institutes of Health, Mammalian Gene Collection (MGC)
ished (1999)
: Daniela S. Gerhard, Ph.D.
of Cancer Genomics
al Cancer Institute / NIH
31 Rm10A07 Bethesda, MD 20892
cgapbs-r@mail.nih.gov
Procurement: Dr. Stefan Hansson
Library Preparation: Michael J. Brownstein (NHGRI) with help
vice from Piero Carninci (RIKEN)
Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
equencing by: Agencourt Bioscience Corporation
distribution: MGC clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL at:
/image.llnl.gov

## Right section

Plate: NDAM612 row: o   column: 14
High quality sequence stop: 287.
FEATURES         Location/Qualifiers
     source      1..1319
                 /organism="Homo sapiens"
                 /mol_type="mRNA"
                 /db_xref="taxon:9606"
                 /clone="IMAGE:30527869"
                 /tissue_type="Human Placenta"
                 /lab_host="DH10B TonA"
                 /clone_lib="NIH MGC 147"
                 /note="Organ: placenta; Vector: pBluescriptR; Site
                 all-XhoI; Site 2: BamH; Oligo-dT primed using prin
                 5'-TTTTTTTTTTTTTTVN-3', size-selected for avera
                 insert size 2.3 kb and normalized to ROT 5. This i
                 primary library enriched for full-length clones ar
                 constructed using the Cap-trapper method (Carninc
                 preparation). Library constructed by M. Brownstei
                 (NIMH/NHGRI, National Institutes of Health). Note:
                 a NIH_MGC library."

ORIGIN

Query Match           12.7%;   Score 175;   DB 14;   Length 1319;
Best Local Similarity 100.0%;  Pred. No. 1.8e-77;
Matches 175;   Conservative   0;   Mismatches   0;   Indels   0;   G

Qy    90  ACAGCCCCCGCCCCCCATGCGCCGCCCGTCGGAGCCAGAGGCGGAGGGGGCGCCGGG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    29  ACAGCCCCCGCCCCCCATGCGCCGCCCGTCGGAGCCAGAGGCGGAGGGGGCGCCGGG

Qy   150  GCCGGGCACCGCCGCCTGGTCCCGCTGGTCCCGCTGGGCCTGGGCCTGGCCGGCGGCCT
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    89  GCCGGGCACCGCCGCCTGGTCCCGCTGGTCCCGCTGGGCCTGGGCCTGGCCGGCGGCCT

Qy   210  CGGCCTCCTGCTCGTGGCCGTGGTCAGTTTGGGAGGCGGGCATCGTGTCCGCCAG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   149  CGGCCTCCTGCTCGTGGCCGTGGTCAGTTTGGGAGGCGGGCATCGTGTCCGCCAG

RESULT 66
AA913913/c
LOCUS       AA913913        338 bp    mRNA    linear    EST 24
DEFINITION  ol35h12.s2 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:1525511 3', mRNA sequence.
ACCESSION   AA913913
VERSION     AA913913.1  GI:3053305
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 338)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contac
            IMAGE Consortium (info@image.llnl.gov) for further informa
            Insert Length: 330   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 332.
FEATURES         Location/Qualifiers
     source      1..338
                 /organism="Homo sapiens"
                 /mol_type="mRNA"
                 /db_xref="taxon:9606"
                 /clone="IMAGE:1525511"
                 /lab_host="DH10B"
                 /clone_lib="Soares_NFL_T_GBC_S1"
                 /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmac
                 a modified polylinker; Site_1: Not I; Site_2: Eco

Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NbHL19W, testis NHT, and B-cell
NCI_CGAP_GCB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo. "

arity    12.7%;    Score 174;    DB 9;    Length 338;
99.7%;    Pred. No. 4.4e-77;    Indels    1;    Gaps    1;
onservative    0;    Mismatches    0;    Indels    1;

CAATCCCTGACCCTTGAGGCCCCCAGTGATCTCGACTCCCCCCTGGCCACAGACC 1138
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAATCCCTGACCCTTGAGGCCCCCAGTGATCTCGACTCCCCCCTGGCCACAGACC 248

GGGCATTGTTGTTCACTGTACTCTGTGGCAAGGATGGGTCCAGAAGACCCCACTTC 1198
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGGCATTGTTGTTCACTGTACTCTGTGGCAAGGATGGGTCCAGAAGACCCCACTTC 188

ACTTAAGAGGGGCTGACCTGGCGCGGCAGGAAGCCAAAGAGACTGGGCTTAGGCCAGG 1258
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ACTTAAGAGGGGCTGACCTGGCGCGGCAGGAAGCCAAAGAGACT-GGCCTAGGCCAGG 129

CCCAAATGTGAGGGGCGAGAAACAAGACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGG 1318
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCAAATGTGAGGGGCGAGAAACAAGACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGG 69

TTAAAACAGATATTATTTTTTATTATTATTATTGTGACAAAATGTTGATAAATGG 1373
||||||||||||||||||||||||||||||||||||||||||||||||||||||
TTAAAACAGATATTATTTTTTATTATTATTATTGTGACAAAATGTTGATAAATGG 14

.97    GI:9703605            196 bp    mRNA    linear    EST 04-AUG-2000
02.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:3221115 3',
sequence.
197
197.1    GI:9703605

sapiens (human)
sapiens
/ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
lia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ases 1 to 196)
GAP http://www.ncbi.nlm.nih.gov/ncicgap.
nal Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Gene Index
lished (1997)
ct: Robert Strausberg, Ph.D.
: cgapbs-r@mail.nih.gov
a Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
nert-Buck, M.D., Ph.D.
Library Preparation: M. Bento Soares, Ph.D., M. Fatima
do, Ph.D.
Library Arrayed by: Greg Lennon, Ph.D.
Sequencing by: Washington University Genome Sequencing Center
e distribution: NCI-CGAP clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL, send email to:
image.llnl.gov
rimer: -40UP from Gibco
quality sequence stop: 175.
Location/Qualifiers
1..196
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3221115"

---

/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_GC6"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modi
polylinker; Site_1: Not I; Site_2: Eco RI; Plasmi
from the normalized library NCI_CGAP_GC4 was prep
ss circles were made in vitro. Following HAP puri
this DNA was used as tracer in a subtractive hybr
reaction. The driver was PCR-amplified cDNAs from
of 5,000 clones made from the same library (clone
1257096-1258631, 1469064-1470983, and 1475592-147
Subtraction by Bento Soares and M. Fatima Bonaldc

ORIGIN

Query Match       12.3%;    Score 169;    DB 10;    Length 196;
Best Local Similarity   100.0%;    Pred. No. 1.4e-74;
Matches 169;    Conservative    0;    Mismatches    0;    Indels    0;    G

Qy  1205  TAAGAGGGGCTGACCTGGCGCGGCAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGAC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   190  TAAGAGGGGCTGACCTGGCGCGGCAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGAC

Qy  1265  CAAATGTGAGGGGCGAGAAACAAGACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGA1
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   130  CAAATGTGAGGGGCGAGAAACAAGACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGA1

Qy  1325  AAAACAGATATTATTTTTATTATTATTGTGACAAAATGTTGATAAATGG 1373
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    70  AAAACAGATATTATTTTTATTATTATTGTGACAAAATGTTGATAAATGG 22

RESULT 68
BI677255/c
LOCUS          BI677255               422 bp    mRNA        linear      EST 1
DEFINITION     id87a02.x1 Human insulinoma Homo sapiens cDNA 3', mRNA se
ACCESSION      BI677255
VERSION        BI677255.1    GI:15630162
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
  ORGANISM     Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutel
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
REFERENCE      1  (bases 1 to 422)
  AUTHORS      Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestn
               Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clift
               Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blist
               Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.
               Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvi
               Williams,T., Jackson,Y. and Bowers,Y.
  TITLE        Endocrine Pancreas Consortium
  JOURNAL      Unpublished (2000)
  COMMENT      Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Ino
               Endocrine Pancreas Consortium
               Harvard University, Howard Hughes Medical Institute
               Dept of Molecular and Cellular Biology, 7 Divinity Ave., C
               MA 02138
               Tel: 617-495-1812
               Fax: 617-495-8557
               Email: dmelton@biohp.harvard.edu
               Library was constructed by Dr. J. Ferrer In vivo mass-exc
               pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washin
               University Genome Sequencing Center For information on ob
               clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.
               Seq primer: -40UP from Gibco
               High quality sequence stop: 416.
               Location/Qualifiers
               1..422
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /tissue_type="insulinoma"
               /lab_host="DH10B (phage-resistant)"
               /clone_lib="Human insulinoma"
FEATURES
     source

/note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
XhoI; Site_2: EcoRI; Constructed with lambda ZAPII
(Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
pBluescript SK- by Dr. H. Inoue following the Washington
University protocol
(http://genome.wustl.edu/est/lambda_protocol.shtml).
Please contact Hiroshi Inoue, MD/PhD for further
information on this library (Metabolism Division, Permutt
Laboratory, Washington University School of Medicine, Box
8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this
is a Washington University Pancreas EST project library."

ORIGIN

Query Match          11.8%;  Score 162;  DB 12;  Length 372;
Best Local Similarity 100.0%; Pred. No. 6e-71;
Matches 162; Conservative 0; Mismatches 0; Indels 0; Ga

Qy    1145  GCATTGTGTTCACTGTACTCTGTGGGCAGGATGGGTCCAGAAGCACCCCACTTCAGC
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     211  GCATTGTGTTCACTGTACTCTGTGGGCAGGATGGGTCCAGAAGCACCCCACTTCAGC

Qy    1205  TAAGAGGGGCTGGACCTGGCGGCAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGAG
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     271  TAAGAGGGGCTGGACCTGGCGGCAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGAG

Qy    1265  CAAATGTGAGGGGCGAGAAACAACAGCAAGCTCTCCCCTTGAG 1306
            |||||||||||||||||||||||||||||||||||||||||||
Db     331  CAAATGTGAGGGGCGAGAAACAAGACAAGCTCTCCCCTTGAG  372

RESULT 70
AQ890280/c
LOCUS       AQ890280                    436 bp    DNA       linear    GSS 10
DEFINITION  HS_3188_B1_F05_MR CIT Approved Human Genomic Sperm Library
            sapiens genomic clone Plate=3188 Col=9 Row=L, genomic surv
            sequence.
ACCESSION   AQ890280
VERSION     AQ890280.1  GI:6346470
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 436)
REFERENCE   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzm
  AUTHORS   Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resge
            BAC end Web Server: http://www.htsc.washington.edu
            Plate: 3188  row: L column: 9
            Seq primer: M13 Reverse
            Class: BAC ends
            High quality sequence stop: 436.
FEATURES             Location/Qualifiers
     source          1..436
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=3188 Col=9 Row=L"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Libr
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clor

---

/note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
XhoI; Site_2: EcoRI; Constructed with lambda ZAPII system
(Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
pBluescript SK- by Dr. H. Inoue following the Washington
University protocol
(http://genome.wustl.edu/est/lambda_protocol.shtml).
Please contact Hiroshi Inoue, MD/PhD for further
information on this library (Metabolism Division, Permutt
Laboratory, Washington University School of Medicine, Box
8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this
is a Washington University Pancreas EST project library."

rity   12.2%;  Score 167;  DB 12;  Length 422;
       100.0%;  Pred. No. 1.7e-73;  Indels 0;  Gaps 0;
nservative 0; Mismatches 0;

       TGCCTGGGCCTGTTCACGTGTTTTCCATCCCACATAAATACAGTATTCCCACTCT 1036
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
       TGCCTGGGCCTGTTCACGTGTTTTCCATCCCACATAAATACAGTATTCCCACTCT  363

       TACAACTCCCCACCGCCCACTCTCCACCTCACTAGCTCCCCAATCCCTGACCCT 1096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
       TACAACTCCCCACCGCCCACTCTCCACCTCACTAGCTCCCCAATCCCTGACCCT  303

       GCCCCCAGTGATCTCGACTCCCCCCCTGGCCACAGACCCCCAG 1143
             |||||||||||||||||||||||||||||||||||||||||||
       GCCCCAGTGATCTCGACTCCCCCCCTGGCCACAGACCCCCAG  256

       372 bp    mRNA      linear    EST 17-SEP-2001
       .y1 Human insulinoma Homo sapiens cDNA 5', mRNA sequence.

       6.1  GI:15630163

apiens (human)
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 372)
D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
ca,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
.,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
.,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
as,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
ns,T., Jackson,Y. and Bowers,Y.
ine Pancreas Consortium
ished (2000)
STs: id87402.x1
: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
ine Pancreas Consortium
d University, Howard Hughes Medical Institute
f Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
38
17-495-1812
17-495-8557
dmelton@biohp.harvard.edu
.y1 was constructed by Dr. J Ferrer In vivo mass-excised to
cript SK- by Dr. H. Inoue DNA sequencing by: Washington
sity Genome Sequencing Center For information on obtaining a
please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
imer: -40RP from Gibco.
Location/Qualifiers
     1..372
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /tissue_type="insulinoma"
     /lab_host="DH10B (phage-resistant)"
     /clone_lib="Human insulinoma"
     /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:

```
Average insert size 1.69 kb. Life Technologies ca
11549-011"

ORIGIN
```

E-Coli DH10B"

```
Query Match        11.8%;   Score 158;  DB 10;   Length 409;
Best Local Similarity   100.0%;   Pred. No. 6.8e-69;
Matches  158;  Conservative  0;  Mismatches  0;  Indels  0;  C
```

```
arity        11.8%;  Score 162;  DB 28;   Length 436;
      99.0%;  Pred. No. 6.2e-71;
onservative  0;  Mismatches  3;  Indels  0;  Gaps  0;
```

```
Qy 1216 GGACCTGCGGCAGGAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGT
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  184 GGACCTGCGGCAGGAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGT
```

```
    CCTGTTCACGTGTTTTCCATCCCACATAATACAGTATTCCCACTCTTATCTTACA 1045
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
    CCTGTTCACGTGTTTTCCATCCCACATAATACAGTATTCCCACTCTTATCTTACA 261
```

```
Qy 1276 GGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTAAAACAG
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  124 GGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTAAAACA(
```

```
    CCCCACCGCCCACTCTCCACCTCACCTAGCTCCCAATCCTGACCCTTTGAGGCCC 1105
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
    CCCCACCGCCCACTCTCCACGTCACTAGCTCCCAATCCTGACCCTTTGAGGCCC 201
```

```
Qy 1336 TATTTTTATTATTGTGACAAAATGTTGATAAATGG 1373
        |||||||||||||||||||||||||||||||||
Db   64 TATTTTTATTATTGTGACAAAATGTTGATAAATGG 27
```

```
    TGATCTTCGACTCCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCT 1165
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
    TGATCTTCGACTCCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCT 141
```

```
    GCAAGGATGGGTCCAGAAGACCCCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCG 1225
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
    GCAAGGATGGGTCCAGAAGACCCCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCG 81
```

```
    GAAGCCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAAC 1285
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
    GAAGCCCAAAGAGACTGGGCATAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAAC 21
```

```
RESULT 72                                        CAAGTCCTCC 1300
R55379                                           ||||||||||
LOCUS       R55379       345 bp   mRNA  linear   CAAGTCCTCC 6
            EST 2?
DEFINITION  yj77a08.r1 Soares breast 2NbHBst Homo sapiens cDNA clone
            IMAGE:154742 5', mRNA sequence.
ACCESSION   R55379
VERSION     R55379.1  GI:824674
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutel
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
REFERENCE   1  (bases 1 to 345)
AUTHORS     Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., a
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. a
            Wilson,R.
TITLE       The WashU-Merck EST Project
JOURNAL     Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 659
            High quality sequence stops: 235 Source: IMAGE Consortium
            This clone is available royalty-free through LLNL ; conta
            IMAGE Consortium (info@image.llnl.gov) for further inform
            Insert Length: 659   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 235.
FEATURES             Location/Qualifiers
     source          1..345
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:55839"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:154742"
                     /sex="Female"
                     /dev_stage="adult"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares breast 2NbHBst"
                     /note="Organ: breast; Vector: pT7T3D (Pharmacia)
                     modified polylinker; Site 1: Not I; Site 2: Eco
                     strand cDNA was primed with a Not I - oligo(dT)
                     TGTTACCAATCTGAAGTGGGAGCGGCGGCCGCCTTTTTTTTTTTTTTTTT
                     double-stranded cDNA was ligated to Eco RI adapt
                     (Pharmacia), digested with Not I and cloned into
                     and Eco RI sites of a modified pT7T3 vector (Pha
                     Library went through one round of normalization
```

```
279    04.x1 NCI_CGAP_Gas4    409 bp   mRNA   linear   EST 27-OCT-1999
04.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2622510 3',
sequence.
279
279.1  GI:6132886

sapiens (human)
yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
lia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ases 1 to 409)

GAP http://www.ncbi.nlm.nih.gov/ncicgap.
nal Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Gene Index
lished (1997)
ct: Robert Strausberg, Ph.D.
: cgapbs-r@mail.nih.gov
e Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
-Buck, M.D., Ph.D.
Library Preparation: Life Technologies, Inc.
Library Arrayed by: Greg Lennon, Ph.D.
Sequencing by: Washington University Genome Sequencing Center
e distribution: NCI-CGAP clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL at:
io.llnl.gov/bbrp/image/image.html
rimer: -40UP from Gibco
quality sequence stop: 405.
       Location/Qualifiers
       1..409
       /organism="Homo sapiens"
       /mol_type="mRNA"
       /db_xref="taxon:9606"
       /clone="IMAGE:2622510"
       /tissue_type="poorly differentiated adenocarcinoma with
       signet ring cell features"
       /lab_host="DH10B"
       /clone_lib="NCI_CGAP_Gas4"
       /note="Organ: stomach; Vector: pCMV-SPORT6; Site 1: SalI;
       Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
```

230. Library constructed by Bento Soares and M.Fatima Bonaldo."

Gruber (Invitrogen). Research Genetics tracking 023. Note: this is a NIH_MGC Library."

ORIGIN

Query Match   9.6%; Score 132; DB 12; Length 785;
Best Local Similarity 99.3%; Pred. No. 1.4e-55;
Matches 302; Conservative 0; Mismatches 1; Indels 1; Ga

Qy 365 GACTAGTTCGGCCTGCGCAGAAGTGCACCTAAAGGCCGGAAAACACGGGCTCGAAGAG
Db 345 GACTAGTTCGGCCTGCGCAGAAGTGCACCTAAAGGCCGGAAAACACGGGCTCGAAGAG

Qy 425 TCGGCAGCCCATTATGAAGTTCATCCACGACCTGACACGGACAGGAGCGCAGGCGAGGTC
Db 405 TCGGCAGCCCATTATGAAGTTCATCCACGACCTGACACGGACAGGAGCGCAGGCGAGGTC

Qy 485 ACGGGACAG-TGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAGCCCCTCTGCGC
Db 465 ACGGGACAGTTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAGCCCCTCTGCGC

Qy 544 AACCGCCAGATCGGGAGTTTATAGTCACCCGGCTCGGGCTCGGGCTCTCACTTGTACTG
Db 525 AACCGCCAGATCGGGAGTTTATAGTCACCCGGCTCGGGCTCGGGCTCTCAGTACTACTG

Qy 604 GTGCACTTTGATGAGGGAAGGTGTGTCACCTGACTTGCTGGACTTGCTGGTGGATGG
Db 585 GTGCACTTTGATGAGGGAAGGTGTGTCACCTGACTTGCTGGACTTGCTGGTGGATGG

Qy 664 CTGG 667
Db 645 CTGG 648

RESULT 74
AI865482/c
LOCUS   AI865482   179 bp   mRNA   linear   EST 30
DEFINITION   wk47a09.x1 NCI_CGAP_Pr22 Homo sapiens cDNA clone IMAGE:241
               mRNA sequence.
ACCESSION   AI865482
VERSION   AI865482.1 GI:5529589
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleo
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 179)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE   National Cancer Institute, Cancer Genome Anatomy Project (
               Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-r@mail.nih.gov
               Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Mi
               Emmert-Buck, M.D., Ph.D.
               cDNA Library Preparation: M. Bento Soares, Ph.D.
               cDNA Library Arrayed by: Greg Lennon, Ph.D.
               DNA Sequencing by: Washington University Genome Sequencin
               Clone distribution: NCI-CGAP clone distribution informati
               found through the I.M.A.G.E. Consortium/LLNL at:
               www-bio.llnl.gov/bbrp/image/image.html
               Seq primer: -40UP from Gibco.
FEATURES   Location/Qualifiers
     source   1..179
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2418520"
                     /sex="male"
                     /tissue_type="normal prostate"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Pr22"
                     /note="Organ: prostate; Vector: pT7T3D-Pac (Pharm
                     with a modified polylinker; 1st strand cDNA was p

rity   10.2%; Score 140; DB 14; Length 345;
       99.0%; Pred. No. 1e-59;
nservative 0; Mismatches 3; Indels 0; Gaps 0;

GTGCACTTTGATGAGGGAAGGCTGTGTCACCTGACTTGCTGGTGGATGG 658
||||||||||||||||||||||||||||||||||||||||||||||||
GTGCACTTTGATGAGGGAAGGCTGTGTCACCTGACTTGCTGGTGGATGG 60

CTGGCCCTGCGCTGCCTGGAGGAATTCTCAGCCACTGCGGGCCCAGTTCCCTCGGGC 718
||||||||||||||||||||||||||||||||||||||||||||||||||||
CTGGCCCTGCGCTGCCTGGAGGAATTCTCAGCCACTGCGGGCCCAGTTCCCTCGGGC 120

CTCCGCCTCTGCCAGGTGTCTGGGCTGTTGGCCCTGCCCGGCCCAGGGTCCTCCCCTGC 778
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTCCGCCTCTGCCAGGTGTCTGGGCTGTTGGCCCTGCCCGGCGCAGGGGTCCTCCCCTGC 180

CGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCTTCCTCACCTACTTCGGAC 838
|||||||||||||||||||||||||||||||||||||||||||||||
CGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCTTCCTCACCTACTTCGGAC 240

CAGGTTCACTGAGGGCCCTGGTCTCCCCACAGTCGTCCCAGGCTGCC 891
||||||||||||||||||||||||||||||||||||||||||
CAGGTTCACTGAGGGCCCTGGTCTCCCCACAGTCGTCCCAGGCTGCC 293

8   785 bp   mRNA   linear   EST 25-SEP-2001
66F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5188168 5',
sequence.

8.1 GI:15754486

piens (human)
piens
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 785)
: http://mgc.nci.nih.gov/.
l Institutes of Health, Mammalian Gene Collection (MGC)
shed (1999)
: Robert Strausberg, Ph.D.
cgapbs-r@mail.nih.gov
Procurement: Life Technologies, Inc.
ibrary Preparation: Life Technologies, Inc.
ibrary Arrayed by: The I.M.A.G.E. Consortium (LLNL)
equencing by: Incyte Genomics, Inc.
distribution: MGC clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov
LLAM1470 row: c column: 17
ality sequence stop: 647.
               Location/Qualifiers
               1..785
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:5188168"
               /lab_host="DH10B"
               /clone_lib="NIH_MGC_116"
               /note="Organ: pooled colon, kidney, stomach; Vector:
               pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
               source anonymous pool of 3 colons, age 26 yo male, 49 yo
               female, 71 yo male colon; 46 yo male kidney, and pool of 2
               stomachs, 62 yo male and 70 yo female. Library is
               oligo-dT primed and directionally cloned (EcoRV site is
               destroyed upon cloning). Average insert size 1.4 kb,
               insert size range 1-3 kb. Library is normalized and
               enriched for full-length clones and was constructed by C.

ORIGIN

```
    Query Match        9.3%;   Score 128;   DB 9;   Length 210;
    Best Local Similarity  100.0%;   Pred. No. 1.2e-53;
    Matches 128;   Conservative   0;   Mismatches   0;   Indels   0;  G

Qy  1246  GGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCC
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   150  GGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCC

Qy  1306  GAATTCCCTGTGGATTTTAAAACAGATATTATTTTATTATTGTGACAAAAT
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    90  GAATTCCCTGTGGATTTTAAAACAGATATTATTTTATTATTGTGACAAAAT

Qy  1366  ATAAAATGG 1373
              |||||||||
Db    30  ATAAAATGG 23

Search completed: April  8, 2004, 23:42:07
Job time : 3973 secs
```

from normal prostate bulk tissue, and was then primed with
a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT7T3 vector. Library is normalized, and was
constructed by Bento Soares and M. Fatima Bonaldo. "

```
arity     9.5%;   Score 130;   DB 9;   Length 179;
100.0%;   Pred. No. 1.1e-54;
onservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

CCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTT 1303
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTT 112

ATTCCCTGTGGATTTTAAAACAGATATTTATTTTTATTATTGTGACAAAATGT 1363
|||||||||||||||||||||||||||||||||||||||||||||||||||||
ATTCCCTGTGGATTTTAAAACAGATATTTATTTTTATTATTGTGACAAAATGT 52

AAAATGG 1373
|||||||
AAAATGG 42
```

```
61
2.61 NCI_CGAP_Br2   210 bp   mRNA   linear   EST 13-AUG-1998
sequence.
61
61.1  GI:3418453

apiens (human)
sapiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 210)
GAP http://www.ncbi.nlm.nih.gov/ncicgap.
al Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Gene Index
ished (1997)
ct: Robert Strausberg, Ph.D.
: cgapbs-r@mail.nih.gov
a Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
-Buck, M.D., Ph.D.
Library Preparation: M. Bento Soares, Ph.D.
Library Arrayed by: Greg Lennon, Ph.D.
Sequencing by: Washington University Genome Sequencing Center
e distribution: NCI-CGAP clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL at:
io.llnl.gov/bbrp/image/image.html
rimer: -40m13 fwd. ET from Amersham
quality sequence stop: 203.
Location/Qualifiers
   1..210
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /clone="IMAGE:1632506"
   /sex="female, pooled"
   /tissue_type="breast"
   /lab_host="DH10B"
   /clone_lib="NCI_CGAP_Br2"
   /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
   polylinker; 1st strand cDNA was prepared from pooled bulk
   breast tumor tissue, and was then primed with a Not I -
   oligo(dT) primer. Double-stranded cDNA was ligated to Eco
   RI adaptors (Pharmacia), digested with Not I and cloned
   into the Not I and Eco RI sites of the modified pT7T3
   vector. This library is the normalized version of
   NCI_CGAP Br1.1. library was constructed by Bento Soares
   and M. Fatima Bonaldo. "
```

GenCore version 5.1.6
opyright (c) 1993 - 2004 Compugen Ltd.

c search, using sw model

il 8, 2004, 19:06:21 ; Search time 654 Seconds
(without alignments)
8918.618 Million cell updates/sec

09-245-198A-3
3
tgtcatgttagacttga.........gacaaatgttgataaatgg 1373

GO_NUC
op 60.0 , Gapext 60.0

3863 seqs, 2124099041 residues

s satisfying chosen parameters: 6747726

th: 0
th: 2000000000

sting first 100 summaries

Geneseq_29Jan04:*
geneseqn1980s:*
geneseqn1990s:*
geneseqn2000s:*
geneseqn2001as:*
geneseqn2001bs:*
geneseqn2002s:*
geneseqn2003as:*
geneseqn2003bs:*
geneseqn2003cs:*
geneseqn2004s:*

the number of results predicted by chance to have a
than or equal to the score of the result being printed,
ed by analysis of the total score distribution.

SUMMARIES

ary ch	Length	DB	ID	Description
0.0	1373	2	AAV18600	Aav18600 Homo sapi
3.6	1306	7	ACC57587	Acc57587 Polynucle
3.6	1306	9	ACC57901	Acc57901 Human TWE
3.6	1306	6	ADC35205	Adc35205 Human CDN
0.8	1364	6	ABK34881	Abk34881 Human CDN
5.4	1353	3	AAA49717	Aaa49717 Human PRO
5.4	1353	6	ABK40255	Abk40255 cDNA enco
5.4	1421	2	AAX56000	Aax56000 Human tum
9.8	1236	4	AAV47613	Aav47613 TNF relat
9.8	1236	4	AAD04350	Aad04350 Human TRE
7.1	1030	2	AAX23424	Aax23424 Human TNR
5.5	898	8	AAS03964	Aas03964 Expressio
9.6	493	8	ACH34013	Ach34013 Human end
3.0	195	6	ABK29540	Abk29540 Colon ade
5.5	282	2	AAT22190	Aat22190 Human gen
4.7	1239	9	ADC97713	Adc97713 Murine FL
4.4	60	6	ABN58848	Abn58848 Human spl
4.4	60	6	ABN41049	Abn41049 Human spl
4.4	60	6	ABN58593	Abn58593 Human spl
4.4	60	6	ABN58591	Abn58591 Human spl
4.4	60	6	ABN58849	Abn58849 Human spl
3.6	50	2	AAX56002	Aax56002 Human tum
3.6	50	3	AAA49732	Aaa49732 Human PRO

No	c	len	ratio	score	n	Accession	Accession	Sp
24		50	3.6	50	6	ABK40292	Abk40292	Oli
25		50	3.6	50	6	ABZ02567	Abz02567	Hun
26		46	3.4	701	2	AAX23425	Aax23425	Mou
27		46	3.4	1168	2	AAV18599	Aav18599	Mu
28		36	2.6	408	7	ABX37032	Abx37032	Bo
29		29	2.1	412	9	ADB56326	Adb56326	To
30		26	1.9	26	6	ABK40356	Abk40356	Re
31		26	1.9	140	7	ABZ78835	Abz78835	Tun
32		26	1.9	140	7	ABZ09382	Abz09382	Hun
33	c	26	1.9	145	7	ABZ78308	Abz78308	Tun
34		26	1.9	145	7	ABZ78428	Abz78428	Tun
35		26	1.9	145	7	ABZ08855	Abz08855	Hun
36		26	1.9	145	7	ABZ08975	Abz08975	Hun
37	c	26	1.9	147	2	AAT19717	Aat19717	Hun
38		26	1.9	153	4	AAI69032	Aai69032	Act
39		26	1.9	281	4	AAS24709	Aas24709	Hun
40		26	1.9	386	6	ABQ60530	Abq60530	Hun
41		26	1.9	391	5	AAH83337	Aah83337	Hun
42	c	26	1.9	425	7	ABX74646	Abx74646	Col
43		26	1.9	439	6	ABL62690	Abl62690	Col
44	c	26	1.9	439	6	ABN96246	Abn96246	Ger
45	c	26	1.9	452	3	AAC06922	Aac06922	Hur
46		26	1.9	469	6	ABV86720	Abv86720	Hur
47		26	1.9	483	6	ABK55088	Abk55088	Hur
48		26	1.9	486	3	AAZ51563	Aaz51563	Hur
49		26	1.9	487	6	ABN94231	Abn94231	Ger
50		26	1.9	498	8	ACH36687	Ach36687	Hur
51		26	1.9	516	3	AAC01272	Aac01272	Hur
52		26	1.9	531	6	ABK55407	Abk55407	Hur
53		26	1.9	540	3	AAC01271	Aac01271	Hur
54		26	1.9	570	3	AAC77898	Aac77898	Hur
55		26	1.9	580	3	AAZ51562	Aaz51562	Hur
56		26	1.9	580	7	ABS56337	Abs56337	Hur
57	c	25	1.8	400	6	ABQ58318	Abq58318	Hur
58		25	1.8	626	5	AAS64426	Aas64426	DN
59		24	1.7	24	2	AAX23451	Aax23451	DN
60	c	24	1.7	24	2	AAX23450	Aax23450	Hur
61	c	24	1.7	40	6	AAX56003	Aax56003	Hur
62		23	1.7	23	6	ABK40355	Abk40355	Pr
63	c	22	1.6	38	2	AAX56004	Aax56004	Hur
64		22	1.6	223	4	AAL18265	Aal18265	Hur
65		22	1.6	263	4	AAL08821	Aal08821	Hu
66		22	1.6	1005	8	ADA02995	Ada02995	Mo
67		22	1.6	1005	9	ADB72733	Adb72733	Mo
68		22	1.6	1005	9	ADC85475	Adc85475	Mo
69		22	1.6	1778	5	AAS84907	Aas84907	DN
70		22	1.6	2942	8	ADA02994	Ada02994	Mo
71		22	1.6	2942	8	ADB72732	Adb72732	Mo
72		22	1.6	2942	9	ADC85474	Adc85474	Mo
73		22	1.6	50295	9	ADA02993	Ada02993	Mo
74		22	1.6	50295	9	ADB72731	Adb72731	Mo
75		22	1.6	50295	9	ADC85473	Adc85473	Mo
76		21	1.5	21	2	AAV18611	Aav18611	Sy
77		21	1.5	21	2	AAV18612	Aav18612	Sy
78		21	1.5	24	2	AAX56006	Aax56006	Hu
79		21	1.5	48	6	AAA49773	Aaa49773	Hu
80	c	21	1.5	48	6	ABK40468	Abk40468	In
81		21	1.5	831	6	ABQ15112	Abq15112	Ol
82	c	21	1.5	831	6	ABQ15113	Abq15113	Ol
83		21	1.5	6182	6	ABL49388	Abl49388	Hu
84	c	21	1.5	7977	6	ABL34088	Abl34088	Hu
85		21	1.5	12025	6	ABL33298	Abl33298	Hu
86		20	1.5	20	2	AAV18603	Aav18603	Sy
87	c	20	1.5	20	2	AAV18610	Aav18610	Sy
88		20	1.5	20	2	AAV18609	Aav18609	Hu
89		20	1.5	20	2	AAV18604	Aav18604	Sy
90		20	1.5	20	2	AAV18602	Aav18602	Sy
91		20	1.5	20	2	AAV18606	Aav18606	Sy
92		20	1.5	20	6	AAV18605	Aav18605	Sy
93		20	1.5	20	6	ABK40354	Abk40354	Fc
94		20	1.5	47	3	AAA49772	Aaa49772	Hu
95		20	1.5	47	2	ABK40467	Abk40467	In
96		20	1.5	196	6	ABL75345	Abl75345	Co

ALIGNMENTS

1.5	455	8	ACH40024	Ach40024 Human foe
1.5	510	8	ACH14153	Ach14153 Human adu
1.5	665	4	AAK93726	Aak93726 Human cDN
1.5	665	4	AAK91714	Aak91714 Human cDN

Query Match   100.0%;   Score 1373;   DB 2;   Length 1373;
Best Local Similarity 100.0%;   Pred. No. 0;
Matches 1373;   Conservative 0;   Mismatches 0;   Indels 0;   G

Qy   1 ATGTCATTGTTAGACTTGAAATTCCGCCCGGCTCCCGTCCCCGATCCCT
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGTCATTGTTAGACTTGAAATTCCGCCCGGCTCCCGTCCCCGATCCCT

Qy   61 TCCCGGATGGGGCGGGGCGGTGAGGCAGGCAGCCCCGCCCCATGGCCGGCCCG
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TCCCGGATGGGGCGGGGCGGTGAGGCAGGCAGCCCCGCCCCATGGCCGGCCCG

Qy 121 AGCCAGAGGCGGAGGGGGCGCGGGGGGAGCGCGGGCACCGCCCTGCTGGTCCGGCT
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 AGCCAGAGGCGGAGGGGGCGCGGGGGGAGCGCGGGCACCGCCCTGCTGGTCCGGCT

Qy 181 CTGGGCCTGGCCTGGCCTGGCCTGCCTCGGCCTCTGCGCCGTGGTCAGTTT
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 CTGGGCCTGGCCTGGCCTGGCCTGCCTCGGCCTCTGCGCCGTGGTCAGTTT

Qy 241 AGCCGGGCATCGTCTGCCCCAGGAGCCTGCCCAGGAGGAGCTGGTGGCAGAGGA
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 AGCCGGGCATCGTCTGCCCCAGGAGCCTGCCCAGGAGGAGCTGGTGGCAGAGGA

Qy 301 CAGGACCCGTCGGAACTGAATCCCAGACAGAAGAAAGCCAGGATCCTGCGCCTTT
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 CAGGACCCGTCGGAACTGAATCCCAGACAGAAGAAAGCCAGGATCCTGCGCCTTT

Qy 361 AACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGCCGGAAAACACGGGCTCC
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 AACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGCCGGAAAACACGGGCTCC

Qy 421 GCGATCGCAGCCATTATGAAGTTCATCCACGACCTGACACAGACGGAGCGCAGG
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 GCGATCGCAGCCATTATGAAGTTCATCCACGACCTGACACAGACGGAGCGCAGG

Qy 481 GTGGACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAGCCCTC
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 GTGGACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAGCCCTC

Qy 541 TACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGGCTTGGCTTCTACTACCTGT
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 TACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGGCTTGGCTTCTACTACCTGT

Qy 601 CAGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGG
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 601 CAGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGG

Qy 661 GTGCTGGCCCTGCGCGCTGCCTGCCAGTTGCCTGGCCTCCACTGCTCCGG
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 661 GTGCTGGCCCTGCGCGCTGCCTGCCAGTTGCCTGGCCTCCACTGCTCCGG

Qy 721 CAGCTCCGCCTCTGCGCCTCCCTGGCCATCTCAAGGCTGCCCCCTTCCTCACCTACTTCG
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 721 CAGCTCCGCCTCTGCGCCTCCCTGGCCATCTCAAGGCTGCCCCCTTCCTCACCTACTTCG

Qy 781 ATCCGCACCCTCCCCTGGGCCATCTCAAGGCTGCCCCCTTCCTCACCTACTTCG
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 781 ATCCGCACCCTCCCCTGGGCCATCTCAAGGCTGCCCCCTTCCTCACCTACTTCG

Qy 841 TTCCAGGTTCACTGAGGGGCCCTGGTCTCCCCACAGTGCGTCCCAGGCTGCCGGCT
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 841 TTCCAGGTTCACTGAGGGGCCCTGGTCTCCCCACAGTGCGTCCCAGGCTGCCGGCT

Qy 901 CGACAGCTCTCTGGCACCCGGTCCCCTGCCCCACCCTCAGCCGCTCTTTGCT
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 901 CGACAGCTCTCTGGCACCCGGTCCCCTGCCCCACCCTCAGCCGCTCTTTGCT

Qy 961 CCTGCCCCTCCCTCTAGAGGCTGCCTGGGCGTGTTTCACGTGTTTCATCCCACA
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 961 CCTGCCCCTCCCTCTAGAGGCTGCCTGGGCGTGTTTCACGTGTTTCATCCCACA

standard; cDNA; 1373 BP.

(first entry)

tumour necrosis factor related ligand (TRELL) gene.

necrosis factor related ligand; tnf; treatment; cancer;
sease; immune system; stimulation; suppression;
on; ds.

Location/Qualifiers
1..852
/*tag= a
/note= "tumour necrosis factor related ligand"

97WO-US013945.

96US-0023541P.
96US-0028515P.
97US-0040820P.

EN INC.
GENEVA FACULTY MEDICINE.

ne Y,   Browning JL;

6619/13.
7525.

a 48-50; 69pp; English.

is that encoding human tumour necrosis factor related ligand
L or active fragments can be included with a carrier in
al compositions to treat cancer, autoimmune diseases or
ses to tissue grafts, or to stimulate or suppress the immune
s useful to screen for TRELL receptors, by labelling with a
abel and screening compositions for binding. Agents
with TRELL-receptor binding can also be screened for, can
nistered, optionally with interferon- gamma, to induce cell
at, suppress or alter immune responses (especially involving
arcinoma cells) involving a signal pathway between TRELL and
. The DNA sequence can be used in gene therapy for TRELL-
rders in mammals (especially humans), e.g. tumours,
nd inflammatory diseases or inherited genetic disorders, by
into cells, and expressing, therapeutically effective amounts
e.g. a virus comprising a gene encoding TRELL. It may also
the preparation of prepare probes for screening
hetic DNAs for TRELL-encoding sequences and for antisense

3 BP; 247 A; 462 C; 394 G; 270 T; 0 U; 0 Other;

Query Match   93.6%; Score 1285; DB 7; Length 1306;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1285; Conservative   0; Mismatches   0; Indels   0; Ga

```
Qy    89 CACAGCCCCCGCCCGCCCATGGCCCGCCCGTTGGAGCCCAGAGGCGGAGCGGGGCGGCGGC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 CACAGCCCCCGCCCGCCCATGGCCCGCCCGTTGGAGCCCAGAGGCGGAGCGGGGCGGCGGC

Qy   149 AGGCGGGCACCGCCCGCTGGTCCGCCCGTTGGAGCCCAGAGACCCACTTCAG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 AGGCGGGCACCGCCCGCTGGTCCGCCCGTTGGGGCCGCATCGCCCCAG

Qy   209 TCGGCCTCTGCTGGCCGTGGTCGTCAGTTTGGGAGCCGGCATGGCCGGTGTCCGCCCAGG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TCGGCCTCTGCTGGCCGTGGTCGTCAGTTTGGGAGCCGGCATGGCCGGTGTCGCCCCAG

Qy   269 CTGCCCAGGAGAGCTGGTGGCAGAGGAGGACCCAGGACCCGTCGGAACTGAATCCC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CTGCCCAGGAGAGCTGGTGGCAGAGGAGGACCCAGGACCCGTCGGAACTGAATCCC

Qy   329 CAGAAGAAAGCCAGGATCCTGCGCTTCCTGAACCGACTAGTTCGGCCTCGCGAGA
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CAGAAGAAAGCCAGGATCCTGCGCTTCCTGAACCGACTAGTTCGGCCTCGCGAGA

Qy   389 CACCTAAAGGCCGGAAAACACGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTT
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CACCTAAAGGCCGGAAAACACGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTT

Qy   449 CACGACCTGACAGGACGGAGCGCAGGAGCAGGTGTGGACGGACAGTGAGTGGCTGG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CACGACCTGACAGGACGGAGCGCAGGAGCAGGTGTGGACGGACAGTGAGTGGCTGG

Qy   509 AAGCCAGAATCAACAGCTCCAGCCCTCTGCGCTACAACCGCCAGATCGGGGAGTTT
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AAGCCAGAATCAACAGCTCCAGCCCTCTGCGCTACAACCGCCAGATCGGGGAGTT

Qy   569 TCACCCGGCTGGGCTCTACTACCTGTACTGTCAGGTGCACTTTGATGAGGGGAGG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TCACCCGGCTGGGCTCTACTACCTGTACTGTCAGGTGCACTTTGATGAGGGGAGG

Qy   629 TCTACCTGAAGCTGACTTGCTGGTGGATGGTGTCTGGCCCTGGCCTGCCTGAGG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TCTACCTGAAGCTGACTTGCTGGTGGATGGTGTCTGGCCCTGGCCTGCCTGAGG

Qy   689 TCTCAGCCACTGCGGCCAGTTCCCTGCGGCCCCAGCTCCGCCTCTGCCAGGTGTCT
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TCTCAGCCACTGCGGCCAGTTCCCTGCGGCCCCAGCTCCGCCTCTGCCAGGTGTCT

Qy   749 TGTTGGCCCTGCCGCCAGGTGCTCCCTGCGGATCCGCACCCTCCCTGGGCCCAT
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TGTTGGCCCTGCCGCCAGGTGCTCCCTGCGGATCCGCACCCTCCCTGGGCCCAT

Qy   809 AGGCTGCCCCTTCCTCACCTACTTCGACTCTTCCAGGTTCACTGAGGGGCCCTG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 AGGCTGCCCCTTCCTCACCTACTTCGACTCTTCCAGGTTCACTGAGGGGCCCTG

Qy   869 CCCCACAGTGTCCCAGGCTGCCGGCTCCCGGGCACCCGGTC
         ||||||||||||||||||||||||||||||||||||||
Db   781 CCCCACAGTGTCCCAGGCTGCCGGCTCCCGGGCACCCGGTC

Qy   929 CTGCCCACCCTCAGCCGCTCTTTGCTCACAAGACCTGCCCCTCCCTAGAGGCTGC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CTGCCCACCCTCAGCCGCTCTTTGCTCACAAGACCTGCCCCTCCCTAGAGGCTGC

Qy   989 GCCTGTTCAGTGTTTTCCATCCACCATAAATACAGTATTCCCACTCTTATCTTAC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 GCCTGTTCAGTGTTTTCCATCCACCATAAATACAGTATTCCCACTCTTATCTTAC

Qy  1049 CCCCACCGCCCACTCTCCACCTACTAGCTCCCAATCCCTGACCCTTTGAGGCC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CCCCACCGCCCACTCTCCACCTACTAGCTCCCAATCCCTGACCCTTTGAGGCC

Qy  1109 GTGATCTCGACTCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGACTC
```

```
ATTCCCACTCTTATCTTACAACTCCCCACGCGCCCACTCTCCACCTTCACTAGCTC 1080
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ATTCCCACTCTTATCTTACAACTCCCCACGCGCCCACTCTCCACCTTCACTAGCTC 1080

TCCCTGACCCTTGAGGCCCCCAGTGATCTCGACTCCCCCTGGCCACAGACCCC 1140
||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCCCCCTGGCCACAGACCCC 1140

CATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAGAGACCCCACTTCAG 1200
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAGAGACCCCACTTCAG 1200

AAGAGGGGCTGGACCTGCGCGCAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGAG 1260
||||||||||||||||||||||||||||||||||||||||||||||||||||||
AAGAGGGGCTGGACCTGCGCGCAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGAG 1260

AAATGTGAGGGGCGAGAAACAAGACAAGCTCCCCTTGAGAATTCCCTGTGGAT 1320
||||||||||||||||||||||||||||||||||||||||||||||||||||||
AAATGTGAGGGGCGAGAAACAAGACAAGCTCCCCTTGAGAATTCCCTGTGGAT 1320

AAACAGATATTATTTTTATTATTGTGACAAAATGTTGATAAATGG 1373
|||||||||||||||||||||||||||||||||||||||||||||
AAACAGATATTATTTTTATTATTGTGACAAAATGTTGATAAATGG 1373
```

ard; DNA; 1306 BP.

first entry)

encoding tumour necrosis factor superfamily member.

tumour necrosis factor; osteopathic; bone; gene; ds.

2.

002WO-US033022.

001US-0329393P.

-JEWISH HOSPITAL.

F, Teitelbaum SL;

46/40.

c comprising a core, and at least one external loop, useful
processes of bone formation or inhibiting bone resorption,
treatments for disease or condition characterized by loss

ge 66-67; 78pp; English.

equence is that of a polynucleotide encoding a non-RANKL
tumour necrosis factor (TNF) superfamily. The invention
naturally-occurring proteins that contain one or more of the
ace loops of RANKL (see ABR42066-70) in combination with a
protein core obtained from a non-RANKL member of the TNF
Also provided are polynucleotides encoding such proteins.
bind to RANK, acting as mimics of RANKL. They can be used to
formation by either inhibiting bone resorption or inducing
thus providing treatment for diseases or conditions
by loss of bone mass

BP; 247 A; 434 C; 368 G; 257 T; 0 U; 0 Other;

XX
SQ   Sequence 1306 BP; 247 A; 434 C; 368 G; 257 T; 0 U; 0 Other;

Query Match   93.6%;   Score 1285;   DB 7;   Length 1306;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1285;   Conservative   0;   Mismatches   0;   Indels   0;   G

```
Qy    89 CACAGCCCCCGCCCCCATGGCCGCCGCCCGTCGGAGCCCAGAGGCCGAGGGGCGGCCGC
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 CACAGCCCCCGCCCCCATGGCCGCCGCCCGTCGGAGCCCAGAGGCCGAGGGGCGGCCGC

Qy   149 AGCCGGGCACCCGCCCTGCTGGTCCGCTGGGCCTGGGCCTGGGCCTGGGCCTGGCCTGGCC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 AGCCGGGCACCCGCCCTGCTGGTCCGCTGGGCCTGGGCCTGGGCCTGGGCCTGGCCTGGCC

Qy   209 TCGGCCTCCTGCTGGCCGTGGTCAGTTTGGGAGCCGGGCATCGCTGCTCCGCCCAC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TCGGCCTCCTGCTGGCCGTGGTCAGTTTGGGAGCCGGGCATCGCTGCTCCGCCCAC

Qy   269 CTGCCCAGGAGGAGCTGGTGGCAGAGAGGACCAGGAGAGGACCCGTCGGAACTGAATCC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CTGCCCAGGAGGAGCTGGTGGCAGAGAGGACCAGGAGAGGACCCGTCGGAACTGAATCC

Qy   329 CAGAAGAAAGCCAGGATCCTGCGCCTTTCCTGACCGACTAGTTCGGCCTCGCAG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CAGAAGAAAGCCAGGATCCTGCGCCTTTCCTGACCGACTAGTTCGGCCTCGCAG

Qy   389 CACCTAAAGGCCGGAAAACACGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGT
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CACCTAAAGGCCGGAAAACACGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGT

Qy   449 CACGACCTGACACGGACCGGAGCCGCAGGCAGGTGTGGACGGGACAGTGAGTGGCTG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CACGACCTGACACGGACCGGAGCGCAGGCAGGTGTGGACGGGACAGTGAGTGGCTG

Qy   509 AAGCCAGAATCAACAGCTCCAGCCCTCCTGCGCTCACAACCGCTACAACGGGGAGTT
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AAGCCAGAATCAACAGCTCCAGCCCTCCTGCGCTCACAACCGCTACAACGGGGAGTT

Qy   569 TCACCCGGGCTGGGCTTCTACTACTGTACTGTCAGGTGCACTTTGATGAGGGGAA
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TCACCCGGGCTGGGCTTCTACTACTGTACTGTCAGGTGCACTTTGATGAGGGGAA

Qy   629 TCTACCTGAAGCTGGACTTGCTGGTGGATGGTGTGCTGGCCCTGCGGCTGCCTGGA
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TCTACCTGAAGCTGGACTTGCTGGTGGATGGTGTGCTGGCCCTGCGGCTGCCTGGA

Qy   689 TCTCAGCCACTGCGGGCCAGTTCCCTCGGGGCCCCAGCTCCGCGGCTCTGCCAGTGTC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TCTCAGCCACTGCGGGCCAGTTCCCTCGGGGCCCCAGCTCCGCGGCTCTGCCAGTGTC

Qy   749 TGTTGGCCCTGCGGCGCCAGGGTCCTCCCCGGGCGGATCCGCACCCTCCCCTGGGCCCA
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TGTTGGCCCTGCGGCGCCAGGGTCCTCCCCGGGCGGATCCGCACCCTCCCCTGGGCCCA

Qy   809 AGGCTGCCCCTTCCTTCCTCACCTACTTCGGACTTCTTCCAGGTTCACTGAGGGGCCCT
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 AGGCTGCCCCTTCCTTCCTCACCTACTTCGGACTTCTTCCAGGTTCACTGAGGGGCCCT

Qy   869 CCCCAGTGTCGTCCCAGGCTGCCGGCTCCCCTGACAGCTCTCGGGCACCCGGT
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 CCCCAGTGTCGTCCCAGGCTGCCGGCTCCCCTGACAGCTCTCGGGCACCCGGT

Qy   929 CTGCCCACCCTCAGCCGCTCTTTGCTCAGACCTGCCCCTCCCTCTAGACGCTG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CTGCCCACCCTCAGCCGCTCTTTGCTCAGACCTGCCCCTCCCTCCTCTAGAGCTG

Qy   989 GCCTGTTCACGTGTTTCCATCCCACATAAATACAGTATTCCCACTCTTATCTTA
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 GCCTGTTCACGTGTTTCCATCCCACATAAATACAGTATTCCCACTCTTATCTTA

Qy  1049 CCCCACCGCCACTCTTCCACCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGC
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
         ||||||||||||||||||||||||||||||
         TCTCGACTCCCCCCTGGCCACAGACCCCCAGGGCATTGTTGTTCACTGTACTCTGTG 1080

Qy      AGGATGGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGCA 1228
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      AGGATGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGCA 1140

Qy      GCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAG 1288
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      GCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAG 1200

Qy      GCTCCTCCCTTGAGAATTCCCTGTGGATTTTAAAAACAGATATTATTTTTATTATT 1348
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      GCTCCTCCCTTGAGAATTCCCTGTGGATTTTAAAAACAGATATTATTTTTATTATT 1260

Qy      TGACAAAATGTTGATAAATGG 1373
         |||||||||||||||||||||
Db      TGACAAAATGTTGATAAATGG 1285
```

dard; cDNA; 1306 BP.

(first entry)

coding sequence.

; tumour necrosis factor; ligand; cytostatic;
tor; osteopathic; gene; ss.

Location/Qualifiers
18..767
/*tag= a
/product= "Human TWEAK"

-A2.

2002WO-US023782.

2001US-0307838P.

N GENOME SCI INC.

Rosen CA;

0659/40.
2315.

ltimeric complex having a first polypeptide member of the
is factor (TNF) ligand family, and a second different member
d family, useful for treating cancer, osteoporosis or an
isease.

Page 367-368; 388pp; English.

sequence is that of a polynucleotide encoding human TWEAK.
n relates to compositions comprising heterotrimeric complexes
crosis factor (TNF) ligand family members, and their use in
n, prevention and treatment of disease. In one embodiment,
imeric complex comprises full-length or extracellular
TWEAK and full-length or extracellular portions of other TNF
y members, preferably VEGI or VEGI-SV. The heterotrimeric
the invention are useful for treating an autoimmune disease,
steoporosis, and particularly for inhibiting cancer cell
n, increasing B cell proliferation, or inducing apoptosis of

```
CC    antibody that binds specifically to endokine alpha polypeptide. Th
CC    present sequence represents a cDNA encoding a tumour necrosis fact
CC    family ligand.
XX
SQ    Sequence 1306 BP; 247 A; 434 C; 368 G; 257 T; 0 U; 0 Other;

Query Match        93.6%;   Score 1285;   DB 9;   Length 1306;
Best Local Similarity  100.0%;   Pred. No. 0;
Matches 1285;  Conservative    0;  Mismatches    0;  Indels    0;   Ga

Qy     89 CACAGCCCCCGCCCACTGGCGCGCCCGTGGAGCCAGAGGCCAGAGAGGCGGAGGGGCGCCGGC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CACAGCCCCCGCCCACTGGCGCGCCCGTGGAGCCAGAGGCCAGAGAGGCGGAGGGGCGCCGGC

Qy    149 AGCCGGGCACCGCCCGTGGTCCGTCCGCTCGCGTGGGCCTGGGCCTGGGCGCTGGCCTGGCCT
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGCCGGGCACCGCCCGTGGTCCGTCCGCTCGCGTGGGCCTGGGCCTGGGCGCTGGCCTGGCCT

Qy    209 TCGGCCTCCTGCTGCCGTGGCTGCAGTTTGGGGAGCGCGGGGCATGCTGCTGTTCCGCCAGC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TCGGCCTCCTGCTGCCGTGGCTGCAGTTTGGGGAGCCGCGGGGCATGCTGCTGTCGCCCAGG

Qy    269 CTGCCCAGGAGAGCTGGTGCAGAGAGGACCAGCAGACCCGTCGGAACTGAATCCCC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CTGCCCAGGAGAGCTGGTGCAGAGAGGACCAGCAGACCCGTCGGAACTGAATCCCC

Qy    329 CAGAAGAAAGCCAGGATCCTGCGCCTTTCCTGAACCGACTAGTTCGGCCTCGCAGA
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CAGAAGAAAGCCAGGATCCTGCGCCTTTCCTGAACCGACTAGTTCGGCCTCGCAGA

Qy    389 CACCTAAAGGCCGGAAAACACGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTTC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CACCTAAAGGCCGGAAAACACGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTTC

Qy    449 CACGACCTGACAGGACGGAGCGAGCCGCAGGTGTGGACGGGACAGTGAGTGGCTGGC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CACGACCTGACAGGACGGAGCGAGCCGCAGGTGTGGACGGGACAGTGAGTGGCTGGC

Qy    509 AAGCCAGAATCAACAGTCCAGCCTCTGCGCTACAACCGCCAGATCGGGGAGTTTA
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AAGCCAGAATCAACAGTCCAGCCTCTGCGCTACAACCGCCAGATCGGGGAGTTTA

Qy    569 TCACCCGGCTGGGCTCTACTACCTGTACAGGTGCACTTTGATGAGGGGAAGC
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TCACCCGGCTGGGCTCTACTACCTGTACAGGTGCACTTTGATGAGGGGAAGC

Qy    629 TCTACCTGAAGCTGGACTTGCTGGTGGATGGTGTCGCCTGGCCTGGCTGGGAGC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TCTACCTGAAGCTGGACTTGCTGGTGGATGGTGTCGCCTGGCCTGGCTGGGAGC

Qy    689 TCTCAGCCACTGCGCGCCAGTTCCCTCGGGCCCCAGCTCCGCCTCTGCCAGTGTCTC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TCTCAGCCACTGCGCGCCAGTTCCCTCGGGCCCCAGCTCCGCCTCTGCCAGTGTCTC

Qy    749 TGTTGGCCCTGCGGCCCAGGGTCTCCCTCGGGATCCGCACCCTCCCCTGGCCCATC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TGTTGGCCCTGCGGCCCAGGGTCTCCCTCGGGATCCGCACCCTCCCCTGGCCCATC

Qy    809 AGGCTGCCCCCTTCCTCACCTACTTCGGACTCTTCACTGAGGGGCCCTGC
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 AGGCTGCCCCCTTCCTCACCTACTTCGGACTCTTCACTGAGGGGCCCTGC

Qy    869 CCCCACAGTGTCCCAGGCTGCCGGCTCCCGCGTCCCCTGACAGTCTCTGGGCACCCGGTC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CCCCACAGTGTCCCAGGCTGCCGGCTCCCGCGTCCCCTGACAGTCTCTGGGCACCCGGTC

Qy    929 CTGCCCACCCTCAGCCGCGTCTTTGCTCACGACCTGCCCCCTCCCTCTAGAGGCTGC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTGCCCACCCTCAGCCGCGTCTTTGCTCACGACCTGCCCCCTCCCTCTAGAGGCTGC

Qy    989 GCCTGTTCACGTGTTTTCCATCCCACCATAAATACAGTATTCCCACTCTTATCTTAC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GCCTGTTCACGTGTTTTCCATCCCACCATAAATACAGTATTCCCACTCTTATCTTAC
```

ACGGCCCACTCTCCACCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCCA 1020

CTTCGACTCCCCCTGGCCACAGACCCCCCAGGGCATTGTGTTCACTGTACTCTGTG 1168
|||||||||||||||||||||||||||||||||||||||||||||||||||||
CTTGACTCCCCCCTGGCCACAGACCCCCCAGGGCATTGTGTTCACTGTACTCTGTG 1080

GGATGGGTCCAGAAGACCCCACTTCAGCGACTTAAGAGGGGCTGGACCTGGCGGGCA 1228
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGATGGGTCCAGAAGACCCCACTTCAGGCACTTAAGAGGGGCTGGACCTGGCGGGCA 1140

CCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGAGGGGCGAGAAACAAG 1288
|||||||||||||||||||||||||||||||||||||||||||||||||
CCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGAGGGGCGAGAAACAAG 1200

CTTCCTCCTTGAGAATTCCCTGTGGATTTTAAAACAGATATTATTTTTATTATT 1348
||||||||||||||||||||||||||||||||||||||||||||||||||||
CTTCCTCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTATTTTTATTATT 1260

GACAAAATGTTGATAAATGG 1373
||||||||||||||||||||
GACAAAATGTTGATAAATGG 1285

ard; cDNA; 1306 BP.

first entry)

oding TNF ligand family member #12.

n; tumour necrosis factor; TNF ligand; endokine alpha;
resorption disorder; osteoporosis; Paget's disease;
fication.

.1.

002US-00218547.

001US-0312542P.
001US-0330761P.

C A.
LI B.

Rosen CA,   Nardelli B;

072/66.
206.

3Q ID NO 23; 145pp; English.

relates to an isolated nucleic acid molecule encoding a
is factor family ligand. A composition comprising the
body or its fragment is used for treating an individual in
ased level of endokine alpha activity. The endokine alpha
resent in a heterotrimeric complex is used for treating an
ving a disorder associated with excessive bone resorption,
osis, Paget's disease or arterial calcification. Treating an
ving a disorder associated with insufficient bone resorption
inistering an endokine alpha antagonist, which is the

alpha gene useful for preparing a composition for treating a
lated with excessive or insufficient bone resorption e.g.,
Paget's disease or arterial calcification.

The proteins are useful in the treatment of various immune defici
and disorders such as viral infections, bacterial infections, fun
infections, autoimmune disorders (e.g. rheumatoid arthritis, mult
sclerosis, autoimmune thyroiditis and diabetes) and allergic reac
and conditions (e.g. asthma). They are also useful for treating
neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's
disease), liver fibrosis, coagulation disorders (e.g. haemophilia
inflammatory disorders (e.g. Crohn's disease) and tumours. They a
useful for tissue regeneration, for wound healing and in the trea
burns, incisions and ulcers. The proteins are also useful for rec
haematopoiesis and for treating myeloid or lymphoid cell deficier
Sequences ABK34863-ABK35454 represent polynucleotides of the inve

```
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
XX
SQ  Sequence 1364 BP; 246 A; 461 C; 394 G; 263 T; 0 U; 0 Other;

Query Match          90.8%;   Score 1247;  DB 6;   Length 1364;
Best Local Similarity 99.9%;  Pred. No. 0;
Matches 1347; Conservative  0; Mismatches  2;  Indels  0;   O
```

```
Qy     25  TCCGCCCGGCCGGCTCCCCTCCCCGATCCCTCCCGGGTCCCCGGATGGGGGGCGG?
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5  TCCGCCCGGCCGGCTCCCCTCCCCGATCCCTCCCGGGTCCCCGGATGGGGGGCGG?

Qy     85  CAGGCACAGCCCCCCGCCCCCCATGGCCGCCCGTCGGAGCCAGGAGGCGGAGGGGGCC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     65  CAGGCACAGCCCCCCGCCCCCCATGGCCGCCCGTCGGAGCCAGGAGGCGGAGGGGGCC

Qy    145  GGGGAGCCGGGCACCGCCCGCTGGTCCGCTGGTCCGCTGGGCCTGGGCCTGGGCGCG?
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    125  GGGGAGCCGGGCACCGCCCGCTGGTCCGCTGGTCCGCTGGGCCTGGGCCTGGGCGCG?

Qy    205  TGCCTCGGCCTCCTGCTCGTCGCCGTGTCAGTTTGGGGAGCCGGGCATCGCTGTCCG?
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    185  TGCCTCGGCCTCCTGCTCGTCGCCGTGTCAGTTTGGGGAGCCGGGCATCGCTGTCCG?

Qy    265  GAGGCTGCCCAGGAGGAGCTGGTGCGCAGAGGAGGACCAGGACCCGTCGGAACTGA?
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    245  GAGGCTGCCCAGGAGGAGCTGGTGCGCAGAGGAGGACCAGGACCCGTCGGAACTGA?

Qy    325  CAGACAGAGAAAGCCAGGATCCTGCGCCTTTCCTGAACCGACTAGTTCGGCCTC?
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    305  CAGACAGAGAAAGCCAGGATCCTGCGCCTTTCCTGAACCGACTAGTTCGGCCTC?

Qy    385  AGTGTCACCCGGGCTGGGCTCGGCTCCTACTACCTGTACTGTCAGCCCATTATG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    365  AGTGTCACCCGGGCTGGGCTCGGCTCCTACTACCTGTACTGTCAGCCCATTATG

Qy    445  CATCCACGACCTGGACAGGACGGAGGCAGGCAGGCAGCGGGACAGTGAGTG
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    425  CATCCACGACCTGGACAGGACGGAGGCAGGCAGGCAGCGGGACAGTGAGTG

Qy    505  GAGGAAGCCAGAATCAACAGCTCAGCCCTCCTCGGCTACAACCGCCAGATCGGGG
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    485  GAGGAAGCCAGAATCAACAGCTCAGCCCTCCTCGGCTACAACCGCCAGATCGGGG

Qy    565  ATAGTCACCCGGGCTGGGCTCGGCTCCTACTACCTGTACTGTCAGGTGCACTTTGATGAGC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    545  ATAGTCACCCGGGCTGGGCTCGGCTCCTACTACCTGTACTGTCAGGTGCACTTTGATGAGC

Qy    625  GCTGTCTCACCTGAAGCTGGACTTGCTGGTGGTGCTCGGCCCTGGCCTGCC
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    605  GCTGTCTCACCTGAAGCTGGACTTGCTGGTGGTGCTCGGCCCTGGCCTGCC

Qy    685  GAATTCTCAGCCACTGCGGCCAGTTCCTCGGCCCCAGTTCCGCCTTCTGCCAGG
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    665  GAATTCTCAGCCACTGCGGCCAGTTCCTCGGCCCCAGTTCCGCCTTCTGCCAGG

Qy    745  GGGCTGTTGGCCCTGCCCAGGGTCTCCGGATCCGCACCCCCCTTGGCC
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db    725  GGGCTGTTGGCCCTGCCCAGGGTCTCCGGATCCGCACCCCCCTTGGCC

Qy    805  CTCAAGGCTGCCCCCTTCCTCACCTACTTCGGACTCTTCCAGGTTCACTGAGGGC?
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

CACCGCCCACTCCCACCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCCA 1108
||||||||||||||||||||||||||||||||||||||||||
CACCGCCCACTCCCACCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCCA 1020

TCTCGACTCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTG 1168
|||||||||||||||||||||||||||||||||||||||||||
TCTCGACTCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTG 1080

AGGATGGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGCA 1228
|||||||||||||||||||||||||||||||||||||||||||
AGGATGGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGCA 1140

GCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGGCGAGAAACAAG 1288
|||||||||||||||||||||||||||||||||||||||||||
GCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGGCGAGAAACAAG 1200

GTCCTCCCTTGAGAAATTCCCTGTGGATTTTTAAAACAGATATTATTTTTTATTATT 1348
|||||||||||||||||||||||||||||||||||||||||||
GCTTCCTCCCTTGAGAAATTCCCTGTGGATTTTTAAAACAGATATTATTTTTTATTATT 1260

TGACAAAATGTTGATAAATGG 1373
|||||||||||||
TGACAAAATGTTGATAAATGG 1285

dard; cDNA; 1364 BP.

(first entry)

coding secreted protein #19.

ed protein; gene; ss; nutritional supplement; haemophilia;
ion; bacterial infection; fungal infection; diabetes; asthma;
isorder; rheumatoid arthritis; multiple sclerosis; tumour;
hyroiditis; allergic reaction; neurodegenerative disease;
disease; Parkinson's disease; liver fibrosis; cancer; ulcer;
disorder; inflammatory disorder; Crohn's disease; incision;
eration; wound healing; burn; haematopoiesis;
deficiency; lymphoid cell deficiency.

A2.

2001WO-US010224.

2000US-0195582P.

TICS INST INC.

ark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;
Graham JR;

9321/23.

e 82; 372pp; English.

l and ninety two polynucleotides derived from a variety of
sources which encode secreted proteins, useful for treating
ciencies and disorders such as autoimmune disorders.

n relates to 592 polynucleotides which have been derived from
human tissue sources and which encode novel secreted
e polynucleotides can be used as probes for the
on and isolation of full length cDNA and genomic DNA. The
des and proteins can also be used as nutritional supplements.

PT   Novel composition to inhibit neoplastic cell growth or for treatin
PT   in mammal comprises polypeptides PRO179, PRO207, PRO219, F
PT   PRO224, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or PRO866.
XX
PS   Claim 20; Fig 3; 172pp; English.
XX
CC   The present sequence is that of cDNA clone DNA30879-1152 (ATCC 20!
CC   encoding human PRO207 (see AAY95338), which shows homology to seve
CC   members of the tumour necrosis factor family, especially human
CC   lymphotoxin (23.4%). The cDNA was identified in a foetal kidney cl
CC   library following identification of an expressed sequence tag wit!
CC   homology to human Apo-2 ligand. A claimed method for inhibiting th
CC   growth of a tumour cell comprises exposing the tumor cell to PRO1'
CC   PRO207, PRO320, PRO219, PRO221, PRO224, PRO328, PRO301, PRO526, PF
CC   PRO356, PRO509 or PRO866 (see AAY95337-49), their agonists or chi
CC   polypeptides incorporating them. The tumour is especially a cance:
CC   selected from breast, ovarian, renal, colorectal, uterine, prosta!
CC   lung, bladder and central nervous system cancer, melanoma and leu!
CC   Nucleic acids encoding PRO179 etc. are used in the recombinant pr(
CC   of the antitumour polypeptides
XX
SQ   Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 U; 0 Other;

Query Match   85.4%; Score 1172; DB 3; Length 1353;
Best Local Similarity   99.8%; Pred. NO. 0;
Matches 1322; Conservative 0; Mismatches   3; Indels   0; G:

Qy   49 CGATCCCTCGGGTCCCGGATGGGGGGGCGGTGAGGCAGGCACAGCCCCCCGCCCC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CGATCCCTCGGGTCCCGGATGGGGGGGCGGTGAGGCAGGCACAGCCCCCCGCCCC

Qy   109 GCCGCCCGTCGGAGCCAGGAGCGGAGGGGGCCCGGGGGAGCCCGGCACCGGCCCT
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GCCGCCCGTCGGAGCCAGGAGCGGAGGGGGCCCGGGGGAGCCCGGCACCGGCCCT

Qy   169 GTCCCGCTCGGCGTGGGCCTGGCGCTGGCGTGGCCTCGGCCTCCTCGCTCGTGGC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTCCCGCTCGGCGTGGGCCTGGCGCTGGCGTGGCCTCGGCCTCCTCGCTCGTGGC

Qy   229 GTCAGTTTGGGGAGCCGGGCATGCGTGTCGCCGCCCAGGAGGCCTGCCCAGGAGCAGGAGCT
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GTCAGTTTGGGGAGCCGGGCATGCGTGTCGCCGCCCAGGAGGCCTGCCCAGGAGGAGCT

Qy   289 GCAGAGGAGGACCAGGACCCGTCGGAACTGGAATCCCCAGACAGAGAAAGCCAGGA
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GCAGAGGAGGACCAGGACCCGTCGGAACTGGAATCCCCAGACAGAGAAAGCCAGGA

Qy   349 GCGCCTTTCCTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAA
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GCGCCTTTCCTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAA

Qy   409 CGGGGCTCGAAGAGCGGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGA
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CGGGGCTCGAAGAGCGGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGA

Qy   469 GCGCAGGCAGGTGTGGACGGACAGTGAGTGGCTGGGGAGGAGGCCAGGAATCAACAG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCGCAGGCAGGTGTGGACGGACAGTGAGTGGCTGGGGAGGAGGCCAGAATCAACAG

GCTGCCCCCTTCCTCCACCTACTACTTCGGACTCTTCCAGGTTCACTGAGGGGCCCTG 844
||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCACAGTCGTCCCAGGCTGCCGGCTCCCCTCGACAGCTCTCTGGGCACCCGGTC 924
||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCGCAGTCGTCGTCCCAGGCTGCCGGCTCCCCTCGACAGCTCTCTGGGCACCCGGTC 904

TGCCCCACCCTCAGCCGCGCTCCTTGCTCCGACACCTGCCCCTCCCTCCTTCTAGAGGCTGC 984
||||||||||||||||||||||||||||||||||||||||||||||||||||
TGCCCCACCCTCAGCCGCGCTCTTTGCTCCAGACCTGCCCCTCCCTCCTTCTAGAGGCTGC 964

CCTGTTCACGTCGTTTTCCATCCCACATAAATACAGTATTCCCACTCTTATCTTAC 1044
||||||||||||||||||||||||||||||||||||||||||||||||||||
CCTGTTCACGTCGTTTTCCATCCCACATAAATACAGTATTCCCACTCTTATCTTAC 1024

CCCCACCGCCCACTCTCCACTTCACTAGCTCCCCAATCCCTGACCCTTTGAGGGCC 1104
||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCCACCGCCCACTCTCCACTTCACTAGCTCCCCAATCCCTGACCCTTTGAGGGCC 1084

TGATCTCGACTCCCCCCCTGGCCACAGACCCCCAGGGGCATTGTTGTTCACTGTACTC 1164
||||||||||||||||||||||||||||||||||||||||||||||||||||
TGATCTCGACTCCCCCCCTGGCCACAGACCCCCAGGGGCATTGTTGTTCACTGTACTC 1144

GCAAGGATGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGAGGGGCTGGACCTGGC 1224
||||||||||||||||||||||||||||||||||||||||||||||||||||
GCAAGGATGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGAGGGGCTGGACCTGGC 1204

GAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAAATGTGAGGGGCGAGAAA 1284
||||||||||||||||||||||||||||||||||||||||||||||||||||
GAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAAATGTGAGGGGCGAGAAA 1264

CAAGCTCTCCCTTGAGAATTCCCTGTGGATTTTAAAACAGATATTATTTTTAT 1344
||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAGCTCTCCCTTGAGAATTCCCTGTGGATTTTAAAACAGATATTATTTTTAT 1324

TTGTGACAAAATGTTGATAAATGG 1373
||||||||||||||||||||||||
TTGTGACAAAATGTTGATAAATGG 1353

lard; cDNA; 1353 BP.

(first entry)

cDNA clone DNA30879-1152.

; antitumour; tumour; therapy; cytostatic; breast cancer;
c; renal cancer; colorectal cancer; uterine cancer;
er; lung cancer; bladder cancer;
is system cancer; melanoma; leukaemia; neoplasm; ss.

2.

Location/Qualifiers
58.   .807
/*tag= a
58. .177
/*tag= b
178. .804
/*tag= c

99WO-US028565.

98US-0113296P.
99WO-US005028.
99US-0130232P.

KW   leukaemia; neuronal disorder; stromal disorder; blastocoelic diso
KW   inflammatory disorder; immune disorder; angiogenic disorder;
KW   gene therapy; cytostatic; neuroprotective; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153486-A1.
XX
PD   26-JUL-2001.
XX
PF   11-FEB-2000; 2000WO-US003565.
XX
PR   08-MAR-1999; 99WO-US005028.
PR   11-MAR-1999; 99US-0123972P.
PR   11-MAY-1999; 99US-0133459P.
PR   02-JUN-1999; 99WO-US012252.
PR   22-JUN-1999; 99US-0140650P.
PR   22-JUN-1999; 99US-0140653P.
PR   20-JUL-1999; 99US-0144758P.
PR   26-JUL-1999; 99US-0145698P.
PR   28-JUL-1999; 99US-0146222P.
PR   17-AUG-1999; 99US-0149395P.
PR   31-AUG-1999; 99US-0151689P.
PR   01-SEP-1999; 99WO-US020111.
PR   15-SEP-1999; 99WO-US021090.
PR   30-NOV-1999; 99WO-US028313.
PR   01-DEC-1999; 99WO-US028301.
PR   01-DEC-1999; 99WO-US028634.
PR   05-JAN-2000; 2000WO-US000219.
XX
PA   (GETH) GENENTECH INC.
XX
PI   Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;
PI   Marsters SA, Pan J, Pitti RM, Roy MA, Smith V, Stone DM;
PI   Watanabe CK, Wood WI;
XX
DR   WPI; 2002-205567/26.
DR   P-PSDB; AAU86129.
XX
PT   Thirty five nucleic acids encoding PRO polypeptides, useful for t
PT   benign or malignant tumors, leukemias and lymphoid malignancies,
PT   inflammatory, angiogenic and immunologic bisorders.
XX
PS   Claim 50; Fig 3; 302pp; English.
XX
CC   The present invention relates to the isolation of novel human PR
CC   polypeptides and the polynucleotide sequences encoding them. The
CC   polypeptides, agonists, antagonists or anti-PRO antibodies are u
CC   treating benign or malignant tumours (e.g. renal, kidney, bladde
CC   breast, etc), leukaemias and lymphoid malignancies, other disord
CC   as neuronal, glial, astrocytal, hypothalamic, glandular, macroph
CC   stromal and blastocoelic disorders, inflammatory, immune and ang
CC   disorders. The polynucleotide sequences are also useful in gene
CC   ABK40254-ABK40288 encode for the human PRO polypeptides of the i
XX
SQ   Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 U; 0 Other;

Query Match   85.4%; Score 1172; DB 6; Length 1353;
Best Local Similarity   99.8%; Pred. No. 0;
Matches 1322; Conservative   0; Mismatches   3; Indels   0;

Qy   49 CGATCCCTCGGGTCCCGGATGGGGGCGGGGGCGGTGAGGCAGGCAGGCACAGCCCCGCCC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CGATCCCTCGGGTCCCGGATGGGGGCGGGGGCGGTGAGGCAGGCAGGCACAGCCCCGCCC

Qy   109 GCCGGCCCGTCGCGAGCCAGAGGCGGAGGGGCGCCGGGGGGAGCCGGGGCACCGCCC
           |||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   61 GCCGGCCCGTCGCGAGCCAGAGGCGGAGGAGCCGGGGGGAGCCGGGGCACCGCCC

Qy   169 GTCCCGCTCGCGCTGGCCGTGGCCTGGCCTCGGCCTCCTGG
           ||||| ||||||||||||||||||||||||| |||||||||
Db   121 GTCCCCGCTCGCGCTGGCCGTGGCCTGGCCGCTGGCCTCCTGG

```
  TCTGCGCTACAACCGCCAGATCGGGAGTTTATAGTCACCCGGCTGGGCTCTAC   588
  |||||||||||||||||||||||||||||||||||||||||||||||||||||
  TCTGCGCTACAACCGCCAGATCGGGAGTTTATAGTCACCCGGCTGGGCTCTAC   540

  GTACTGTCAGTGCACTTTGATGAGGGAAGGCTGTCTCACCTGAAGCTGGACTTG   648
  |||||||||||||||||||||||||||||||||||||||||||||||||||||
  GTACTGTCAGTGCACTTTGATGAGGGAAGGCTGTCTCACCTGAAGCTGGACTTG   600

  TGGATGGTGTGCTGGCCCTGCGCTGCCTGGAGGAATTCTCAGCCACTGCGGCCAGT   708
  |||||||||||||||||||||||||||||||||||||||||||||||||| ||||
  TGGATGGTGTGCTGGCCCTGCGCTGCCTGGAGGAATTCTCAGCCACTGCGGCGAGT   660

  TCGGGCCCCAGCTCCGCCTCTTGCCAGTGTCTGGCGTGTTGGCCCTTGCGCGCCAGGG   768
  ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
  TCGGGCCCCAGCTCCGCCTCTTGCCAGTGTCTGGCGTGTTGGCCCTGCGCGCCAGGG   720

  CCCTGCGGATCCGACACCCTCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACC   828
  |||||||||||||||||||||||||||||||||||||||||||||||||||||
  CCCTGCGGATCCGACACCCTCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACC   780

  TCGGACTCTTCCAGGTTCACTGAGGGGGCCCTGGTCTCCCCACAGTCGTCCCAGGCT   888
  ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
  TCGGACTCTTCCAGGTTCACTGAGGGGGCCCTGGTCTCCCCGCAGTCGTCCCAGGCT   840

  GCTCCCCTCGACAGCTCTCTCGGGCACCCGGTCCCCTGCCCCACCCTCAGCCGCGCT   948
  |||||||||||||||||||| ||||||||||||||||||||||||||||||||||
  GCTCCCCTCGACAGCTCTCTGGGCACCCGGTCCCCTGCCCCACCCTCAGCCGCGCT   900

  GCTTCCAGACCTGCCCCCTCCCCTCCCTCTAGAGCGCTGCCTGGCCGTCCGTGTTTCCA  1008
  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  GCTTCCAGACCTGCCCCCTCCCCTCCCTCTAGAGCGCTGCCTGGCCGTCCGTGTTTTCCA  960

  ACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCCACTCTCCA   1068
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
  ACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCCACTCTCCA   1020

  CACTAGCTCCCCAATCCCTGACCCTTGAGGCCCCCAGTGATCTCGACTCCCCCTG   1128
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
  CACTAGCTCCCCAATCCCTGACCCTTGAGGCCCCCAGTGATCTCGACTCCCCCTG   1080

  ACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAG   1188
  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  ACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAG   1140

  CACTTCAGGCACTAAGAGAGGGCTGACCTGGCCGGCGAGAAGCCAAAGAGACTGGGC   1248
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  CACTTCAGGCACTAAGAGAGGGCTGACCTGGCCGGCGAGAAGCCAAAGAGACTGGGC   1200

  GGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAA   1308
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  GGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAA   1260

  CCTGTGTGGATTTTTAAAACAGATATTATTTTTATTATTCTGACAAAATGTTGATA   1368
  |||||||||||||||||||||||||||||||||||||||||||||||||||||||
  CCTGTGTGGATTTTTAAAACAGATATTATTTTTATTATTCTGACAAAATGTTGATA   1320
```

GG 1373
||
GG 1325

ndard; cDNA; 1353 BP.

(first entry)

g human PRO207 polypeptide.

benign tumour; malignant tumour; lymphoid malignancy;

Db   1261 TTCCCTGTGGATTTTTAAAACAGATATTATTTTTATTATTTGTGACAAAATGTTC
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   1369 AATCG 1373
           |||||
Db   1321 AATCG 1325

RESULT 8
AAX56000
ID   AAX56000 standard; DNA; 1421 BP.
XX
AC   AAX56000;
XX
DT   15-JUL-1999 (first entry)
XX
DE   Human tumour necrosis factor Apo-3 ligand polynucleotide sequence
XX
KW   Human; tumour necrosis factor; Apo-3 ligand; lymphotoxin; apoptos
KW   NF-kappaB-dependent transcription; JNK/SAPK-dependent response; c
KW   ss.
XX
OS   Homo sapiens.
XX
FH   Location/Qualifiers
FT   Key
FT   CDS   92..841
FT   /*tag= a
FT   /product= "Apo-3 ligand"
XX
PN   WO9919490-A1.
XX
PD   22-APR-1999.
XX
PF   09-OCT-1998;   98WO-US021407.
XX
PR   10-OCT-1997;   97US-0062037P.
PR   17-DEC-1997;   97US-0069862P.
XX
PA   (GETH) GENENTECH INC.
XX
PI   Ashkenazi AJ, Marsters SA, Pitti R;
XX
DR   WPI; 1999-287982/24.
DR   P-PSDB; AAY09369.
XX
PT   New human Apo3- ligand (a tumor necrosis factor) homologue.
XX
PS   Claim 18; Fig 1; 74pp; English.
XX
CC   The present sequence encodes a human tumour necrosis factor (TNF)
CC   lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has
CC   cytostatic activity. Apo-3 ligand can be used to induce apoptosi
CC   mammalian cancer cells, to induce NF-kappaB-dependent transcript
CC   to induce JNK/SAPK-dependent responses in mammalian cells
XX
SQ   Sequence 1421 BP; 281 A; 464 C; 404 G; 272 T; 0 U; 0 Other;

Query Match   85.4%; Score 1172; DB 2; Length 1421;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1322; Conservative   0; Mismatches   3; Indels   0; (

Qy   49 CGATCCCCTCGGGTCCCGGGATGGGGGCGGTGAGGCGGTGAGGCGGTGAGGCAGCCCCCCGCCCC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   35 CGATCCCCTCGGGTCCCGGGATGGGGGCGGTGAGGCGGTGAGGCGGTGAGGCAGCCCCCCGCCCC

Qy   109 GCCGCCCGTCGGAGCCAGAGGCCGGAGGGGGCGCCCGGGGGGAGCCGGGCACCGCGCCC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   95 GCCGCCCGTCGGAGCCAGAGGCCGGAGGGGGCGCCCGGGGGGAGCCGGGCACCGCGCCC

Qy   169 GTCCCGCTCGCGCTGGGCGTCTGGGCCTGCCTCGGCCTCGCCGCTCCTGTGCTCTG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   155 GTCCCGCTCGCGCTGGGCGTCTGGGCCTGCCTCGGCCTCGCCGCTCCTGTGCTCTG

Sequence alignment (query/database):

TTTGGGGAGCCGGGCATCGCTGTCCGCCCAGGAGGCCTGCCCAGGAGGAGGAGCTGCGTG 288
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TTTGGGGAGCCGGGCATCGCTGTCCGCCCAGGAGGCCTGCCCAGGAGGAGGAGCTGCGTG 240

GGAGGACCAGGACACCGTCGGAACTGAATCCCCAGACAGAAGAAAGCCAGGATCCT 348
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGAGGACCAGGACACCGTCGGAACTGAATCCCCAGACAGAACAGAGAAGAAAGCCAGGATCCT 300

TTTCCTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAACA 408
||||||||||||||||||||||||||||||||||||||||||||||||||||||
TTTCCTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAACA 360

TCGAAGAGGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGACAGGACCGA 468
||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCGAAGAGCGATGCAGCCCATTATGAAGTTCATCCACGACCTGACAGGACCGA 420

GGCAGGTGTGGACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCC 528
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGCAGGTGTGGACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCC 480

TCTGCGCTACAACCGACCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTAC 588
||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCTGCGCTACAACCGACCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTAC 540

GTACTGTCAGGTGCACTTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGACTTG 648
|||||||||||||||||||||||||||||||||||||||||||||||||||||
GTACTGTCAGGTGCACTTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGACTTG 600

GGGATGGTGTGTCTGGCCCTGCGCTCGCCTGGAGGAATTCTCAGCCACTGCGGCCAGT 708
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGGATGGTGTGTCTGGCCCTGCGCTCGCCTGGAGGAATTCTCAGCCACTGCGGCGAGT 660

TCGGGCCCCAGCTCCGCCTCTGCCAGTGTCTGTTGGCCCTGCGCGCCAGGG 768
||||||||||||||||||||||||||||||||||||||||||||||||||
TCGGGCCCCAGCTCCGCCTCTGCCAGTGTCTGTTGGCCCTGCGCGCCAGGG 720

CCCTGCCGGATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCCTTCCTCACC 828
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCTGCCGGATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCCTTCCTCACC 780

TCGGACTCTTCCAGGTTCACTGAGGGGCCCTGTCCCCCACAGTGTCCCAGGCT 888
|||||||||||||||||||||||||||||||||||||||||||||||||||
TCGGACTCTTCCAGGTTCACTGAGGGGCCCTGTCCCCCGCAGTGTCCCAGGCT 840

GCTCCCCTCGACAGCTCTCTGGCACCCGGTCCCCCTCTGCCCCACCCCTCAGCCGCT 948
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCTCCCCTCGACAGCTCTCTGGCACCCGGTCCCCCTCTGCCCCACCCCTCAGCCGCT 900

GCTCCAGACCTGCCCCTCCCTCCAGAGGCTGCCTGGGCCTGTTCACGTGTTTTCCA 1008
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCTCCAGACCTGCCCCTCCCTCTAGAGGCTGCCTGGGCCTGTTCACGTGTTTTCCA 960

ACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCACTCTCCA 1068
|||||||||||||||||||||||||||||||||||||||||||||||||||||
ACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCCACTCTCCA 1020

ACTAGCTCCCAATCCCTGACCCTTTGAGGCCCCAGTGATCTCGACTCCCCCCTG 1128
||||||||||||||||||||||||||||||||||||||||||||||||||||
ACTAGCTCCCAATCCCTGACCCTTTGAGGCCCCAGTGATCTCGACTCCCCCCTG 1080

ACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAG 1188
|||||||||||||||||||||||||||||||||||||||||||||||||||||
ACAGACCCCCAGGTCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAG 1140

CCACTTCAGGCACTAAGAGGGGCTGGACCTTGGCGCGGCAGAAGCCAAAGAGACTCGGGC 1248
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCACTTCAGGCACTAAGAGGGGCTGGACCTTGGCGCGGCAGGAGCCAAAGAGACTGGGC 1200

GGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTTGAGAA 1308
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTTGAGAA 1260

CCTGTGTGGATTTTTAAAACAGATATTATTTTTATTATTTGTGACAAAATGTTGATA 1368

```
Db   1295 TTCCCGTGGATTTTAAAACAGATATTTTATTATTGTGACAAAATGTT
          |||||||||||||||||||||||||||||||||||||||||||||||
Qy   1369 AATGG 1373
          |||||
Db   1355 AATGG 1359
```

RESULT 9
AAV47613
ID   AAV47613 standard; cDNA; 1236 BP.
XX
AC   AAV47613;
XX
DT   27-OCT-1998   (first entry)
XX
DE   TNF related endothelium proliferative agent gene.
XX
KW   ss; TNF; endothelium proliferative agent; TREPA; wound healing; (
KW   tissue grafting; vascularisation; apoptosis; autoimmune; birth c(
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..750
FT   /*tag= a
FT   /product= "TREPA"
XX
PN   WO9835061-A2.
XX
PD   13-AUG-1998.
XX
PF   12-FEB-1998;   98WO-US002859.
XX
PR   12-FEB-1997;   97US-00798692.
PR   10-FEB-1998;   98US-00021706.
XX
PA   (ABBO) ABBOTT LAB.
XX
PI   Wiley SR;
XX
DR   WPI; 1998-447255/38.
DR   P-PSDB; AAW29745.
XX
PT   Detecting nucleic acid encoding TREPA - useful for diagnosis and
PT   treatment of autoimmune disease, tumours and inflammation.
XX
PS   Claim 11; Page 123-4; 142pp; English.
XX
CC   The TNF-related endothelium proliferative agent (TREPA), or its
CC   activators or agonists, are used to treat a deficit of TREPA, e.
CC   promote wound healing or tissue grafting, by promoting vascular:
CC   also to induce apoptosis for treating cancer and eliminating aut
CC   T cells, as an adjunct to cancer chemotherapy or antiviral trea:
CC   TREPA peptides can also be used to target cytotoxic agents or f:
CC   affinity isolation of the corresponding receptor, the nucleic a:
CC   which can be used to transform tumour cells to render them more
CC   responsive to TREPA and to screen for TREPA mimics. Ribozymes, :
CC   RNA, antibodies or peptides, are used to treat TREPA-associate:
CC   diseases, e.g. tumours and metastases (by inhibiting vascularis.
CC   inflammation or a wide range of autoimmune conditions, conditio
CC   involving abnormal stimulation of epithelial cells (e.g.
CC   atherosclerosis), for birth control (inhibiting ovulation and p
CC   formation) or other angiogenic conditions (e.g. ulcers)
XX
SQ   Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 U; 0 Other;

```
Qy   129 GCGGAGGGGGGCGCCCGGGGGGAGCCGGCACCGCCCCTGGTCCCGCTCGCGCT
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Query Match   69.8%;   Score 958;   DB 2;   Length 1236;
Best Local Similarity   99.6%;   Pred. No. 0;
Matches 1208;   Conservative   0;   Mismatches   5;   Indels   0;

```
Qy GTTTGGGAGCCGGCATCGCTGTCCGCCCAGGAGCCTGGTG 288
   ||||||||||||||||||||||||||||||||||||||||
Db GTTTGGGAGCCGGCATCGCTGTCCGCCCAGGAGCCTGGTG 274

Qy AGGAGGACCAGGACCCGTCGGAACTGAATCCCAGACAGAAGAAGCCAGGATCCT 348
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db AGGAGGACCAGGACCCGTCGGAACTGAATCCCAGACAGAAGAAGCCAGGATCCT 334

Qy CTTTCCTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAACA 408
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db CTTTCCTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAACA 394

Qy CTCGAAGAGCGATCGCAGCCATTATGAGAGTTCATCCACGACCTGGACAGGACCGA 468
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db CTCGAAGAGCGATCGCAGCCATTATGAGAGTTCATCCACGACCTGGACAGGACCGGA 454

Qy AGGCAGGTGTGGACGGGACAGTGAGTGGGCTGGGAGGAAGCCAGAATCAACAGCTCC 528
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db AGGCAGGTGTGGACGGGACAGTGAGTGGGCTGGGAGGAAGCCAGAATCAACAGCTCC 514

Qy CTCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGCTGGGCTCTAC 588
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db CTCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGCTGGGCTCTAC 574

Qy CTGTACTGTCAGGTGCACTTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGGACTTG 648
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db CTGTACTGTCAGGTGCACTTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGGACTTG 634

Qy GTGGATGGTGTGCTGGCCCTGCCGCTGCCTGAGGGAATTCTCAGCCACTGCCGGCCAGT 708
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db GTGGATGGTGTGCTGGCCCTGCCGCTGCCTGAGGGAATTCTCAGCCACTGCCGCGAGT 694

Qy CTCGGGCCCCAGCTCCCGCCTCTGCCAGGTGTCTGGGCTGTTGGCCCTGCGGCCAGGG 768
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db CTCGGGCCCCAGCTCCCGCCTCTGCCAGGTGTCTGGGCTGTTGGCCCTGCGGCCAGGG 754

Qy TCCCTGCGGATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCTTCCTCACC 828
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db TCCCTGCGGATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCTTCCTCACC 814

Qy TTCGGACTCTTCCAGGTTCACTGAGGGGCCCTGGTCTCCCCACAGTCGTCCCAGGCT 888
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db TTCGGACTCTTCCAGGTTCACTGAGGGGCCCTGGTCTCCCCGAGTCGTCCCAGGCT 874

Qy GGCTCCCCTGACCAGCTCTCTGGGCACCCGGTCCCCGTCCCCTCGCCCACCCTCAGCCGCT 948
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db GGCTCCCCTGCGACAGCTCTCTGGGCACCCGGTCCCCGTCCCCTCGCCCACCCTCAGCCGCT 934

Qy TGCTCCAGACCTGCCCTCCTCCTCTAGAGGCTGCCTGGGCTGTTCACGTGTTTTCCA 1008
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db TGCTCCAGACCTGCCCTCCTCCTCTAGAGGCTGCCTGGGCTGTTCACGTGTTTTCCA 994

Qy CACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCCACTCTCCA 1068
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db CACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCCACTCTCCA 1054

Qy CACTAGCTCCCAATCCTGACCCTTTGAGCCCCCAGTGATCTCGACTCCCCCCTG 1128
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db CACTAGCTCCCAATCCTGACCCTTTGAGCCCCCAGTGATCTCGACTCCCCCCTG 1114

Qy CACAGACCCCCAGGCATTGTTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAG 1188
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db CACAGACCCCCAGGCATTGTTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAG 1174

Qy CCCACTTCAGCGACTCAAGAGGGGGCTGACCTGGACCTGGCGGCAGGAAGCCAAAGAGACTGGGC 1248
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db CCCACTTCAGCGACTCAAGAGGGGGCTGACCTGGACCTGGCGGCAGGAAGCCAAAGAGACTGGGC 1234

Qy AGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAAACAAGACAGCTCTCCCCTTGAGAA 1308
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db AGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAAACAAGACAGCTCTCCCCTTGAGAA 1294

Qy CCCTGTGTGGATTTTTAAAACAGATATTTTATTATTATTGTGACAAAATGTTGATA 1368
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy    1269  TGTGAGGGGCGAGAAACAAGAACAAGAGCTCTCCCTTGAGAATTCCCTGTGGATTTTT
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1164  TGTGAGGGGCGAGAAACAAGAACAAGAGCTCTCCCTTGAGAATTCCCTGTGGATTTTT

Qy    1329  CAGATATTATTT  1341
            ||||||||||||
Db    1224  CAGATATTATTT  1236
```

RESULT 10
AAD04350
ID   AAD04350 standard; cDNA; 1236 BP.
XX
AC   AAD04350;
XX
DT   04-JUL-2001 (first entry)
XX
DE   Human TREPA (TNF related endothelium proliferative agent) cDNA.
XX
KW   Human; tumour necrosis factor; TNF; angiogenesis; wound healing;
KW   TNF related endothelium proliferative agent; tumour; metastasis;
KW   grafting; vulnerary; ss
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..750
FT   /*tag= a
FT   /product= "Human TREPA (TNF related endothelium
FT   proliferative agent)"
XX
PN   US6207642-B1.
XX
PD   27-MAR-2001.
XX
PF   26-JUN-1998;   98US-00105343.
XX
PR   12-FEB-1997;   97US-00798692.
PR   10-FEB-1998;   98US-0021706.
XX
PA   (ABBO) ABBOTT LAB.
XX
PI   Wiley SR;
XX
DR   WPI; 2001-280760/29.
DR   P-PSDB; AAE00891.
XX
PT   Inducing angiogenesis in mammal at desired sites for promoting wo
PT   healing, by administering soluble fragment of extracellular doma
PT   tumor necrosis factor related endothelium proliferative agent pro
XX
PS   Example 2; Col 73-74; 53pp; English.
XX
CC   The present invention relates to extracellular signal molecules,
CC   particularly members of tumour necrosis factor (TNF) family molec
CC   designated as TREPA (TNF related endothelium proliferative agent)
CC   Soluble biologically active TREPA are used to treat TREPA-associa
CC   diseases, tumours or metastases. TREPA is used for inducing angic
CC   in human for promoting wound healing and for vascularising grafts
CC   for successful grafting and to promote tissue grafts. The present
CC   sequence is a cDNA clone ID #690050 encoding human TREPA
XX
SQ   Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 U; 0 Other;

Query Match   69.8%;   Score 958;   DB 4;   Length 1236;
Best Local Similarity   99.6%;   Pred. No. 0;
Matches 1208;   Conservative   0;   Mismatches   5;   Indels   0; (

```
Qy    129  GCGGAGGGGGCGCCGGGGGGGAGCGGGGACCGCGGGCACCGCCCCTGGTCCCGCGCTGG
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     24  GCGGAGGGGGCGCCGGGGGGGAGCGGGGACCGCGGGCACCGCCCCTGGTCCCGCGCTGG
```

```
           SGGGGGGCCGCGGGGAGCGGGGCACCGCCCGCTGCTGTCCGCTGCGCTCGCGCTGGGCCT    83
           TGGCGCTGGCCTGCCTGCCTGCCTGCCGTGGTCAGTTGGGGAGCCGGGC    248
           |||||||||||||||||||||||||||||||||||||||||||||||||
           TGGCGCTGGCCTGCCTGCCTGCCTGCCGTGGTCAGTTTGGGGAGCCGGGC    143

           TGTCCGCCCAGGAGCCTGCCCAGGAGGAGCTGGTGGCAGAGGAGGACCAGGACCC    308
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
           TGTTCCGCCCAGGAGCCTGCCCAGGAGGAGCTGGTGGCAGAGGAGGACCAGGACCC    203

           AACTGAATCCCCAGACAGAGAAGAAAGCCAGGATCCTCGGCCTTTCCTGAACCGACT    368
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
           AACTGAATCCCCAGACAGAGAAGAAAGCCAGGATCCTCGCGCCTTTCCTGAACCGACT    263

           GGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAACACGGGCTCGAAGAGCGGATCGC    428
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
           GGCCTCGAAGAAGTGCACCTAAAGGCCGGAAAACACGGGCTCGAAGAGCGGATCGC    323

           ATTATGAAGTCATCCACAGACCTGACACGAGACGGAGCGCAGGCAGGTGTGGACGG    488
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
           ATTATGAAGTCATCCACAGACCTGACAGACAGGAGCGCAGGCAGGTGTGGACGG    383

           TGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAGCCCTCTGCGCTACAACCG    548
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
           TGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAGCCCTCTGCGCTACAACCG    443

           ATCGGGGGAGTTTATAGTCACCCGGGCTGGGCTCTACTACCTGTACTGTCAGGTGCA    608
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
           ATCGGGGGAGTTTATAGTCACCCGGGCTGGGCTCTACTACCTGTACTGTCAGGTGCA    503

           GATGAGGGGAAGGCTGTCTACCTGAAGCTGACTTGCTGGTGATGGTGTGTGCTGGC    668
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
           GATGAGGGGAAGGCTGTCTACCTGAAGCTGACTTGCTGGTGATGGTGTGTGCTGGC    563

           CGCTGCCTGGAGGAATTCTCAGCCACTGCGCGCCAGTTCCCTGCGGGCCCCAGCTCCG    728
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
           CGCTGCCTGGAGGAATTCTCAGCCACTGCGGCGGCGAGTTCCCTGCGGGCCCCAGCTCCG    623

           TGCCAGTGTGTGTTGGGCGTTGGCCCTGGCCCTGCGGGTCCTCCCTGCGGGATCCGCAC    788
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
           TGCCAGTGTGTGTTGGGCGTTGGCCCTGGCCCTGCGGGTCCTCCCTGCGGGATCCGCAC    683

           CCCTGGGCCCATCTCAAGGCTGCCCCCCTTCCTCACCTACTTCGGACTCTTCCAGGT    848
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
           CCCTGGGCCCATCTCAAGGCTGCCCCCCTTCCTCACCTACTTCGGACTCTTCCAGGT    743

           TGAGGGGCCCTGGTCTCCCCACAGTCGTCCCCAGTGCTGCCGGGCTCCCCTGACAGCT    908
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
           TGAGGGGCCCTGGTCTCCCCCGCAGTCGTCCCCAGTGCTGCCGGGCTCCCCTGACAGCT    803

           GGGCACCCGGTCCCCTCGCCCCACCCTCAGCCGCTCTTGCTCCAGACCTGCCCC    968
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
           GGGCACCCGGTCCCCTCGCCCCACCCTCAGCCGCTCTTGCTCCAGACCTGCCCC    863

           TCTAGAGGCTGCCTGGGCCTGTTCACGTGTTTTCCATCCACATAAATACAGTATT    1028
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
           TCTAGAGGCTGCCTGGGCCTGTTCACGTGTTTTCCATCCACATAAATACAGTATT    923

           CTCTTATCTTACAACTCCCCACCGCCACTCCACCTCCACCTCAGTCCCCAATCC    1088
           |||||||||||||    |||||||||||||||||||||||||||||||||||||
           CTCTTATCTTACACACAACCCCACCGCCACTCCACCTCCACCTCAGTCCCCAATCC    983

           CCCTTTGAGGCCCCCAGTGATCTGACTCCCCCCTGGCCACAGACCCCCAGGGCAT    1148
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
           CCCTTTGAGGCCCCCAGTGATCTGACTCCCCCCTGGCCACAGACCCCCAGGGCAT    1043

           TTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCCCACTTCAGGCACTAAG    1208
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
           TTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCCCACTTCAGGCACTAAG    1103

           GCTGGACCTGGCGGCAGGAAGCCAAAGAGACTGGGCTTAGGCCAGGAGTTCCCAAA    1268
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
           GCTGGACCTGGCGGCAGGAAGCCAAAGAGACTGGGCTTAGGCCAGGAGTTCCCAAA    1163
```

```
Db   1164 TGTGAGGGGCGAGAAACAAGACAAGCAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTT
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   1329 CAGATATTATTTT 1341
          |||||||||||||
Db   1224 CAGATATTATTTT 1236
```

RESULT 11
AAX23424
ID   AAX23424 standard; DNA; 1030 BP.
XX
AC   AAX23424;
XX
DT   18-JUN-1999 (first entry)
XX
DE   Human TNRL3 DNA.
XX
KW   Tumour necrosis factor receptor; signal transducer molecule; TNF;
KW   developmental abnormality; gestational abnormality; prostate ca
KW   APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy;
KW   cytoplasmic domain; immunogen; antibody preparation; breast/carc
KW   apoptosis; human; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..627
FT   /*tag= a
FT   /product= "TNRL3"
XX
PN   WO9911791-A2.
XX
PD   11-MAR-1999.
XX
PF   04-SEP-1998; 98WO-US018393.
XX
PR   05-SEP-1997; 97US-00924634.
XX
PA   (UNIW) UNIV WASHINGTON.
XX
PI   Chaudhary PM;
XX
DR   WPI; 1999-205191/17.
DR   P-PSDB; AAW93590.
XX
PT   New Tumor Necrosis Factor family receptor polypeptides and ligan
PT   useful for diagnosis and treatment of prostate cancer and develo
PT   or gestational abnormalities.
XX
PS   Example VII; Fig 13A; 156pp; English.
XX
CC   This invention describes isolated Tumor Necrosis Factor (TNF) fa
CC   receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC   fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and T
CC   their active fragments. APO4 is useful for diagnosing prostate c
CC   determining levels of APO4 in an individual. Prostate cancer can
CC   treated using APO4 selective binding agents linked to a therapeu
CC   moiety. APO4 polypeptides are also useful for identifying select
CC   binding agents, useful in diagnosis/treatment of disease by bind
CC   agents to the polypeptide/active fragment which is extracellular
CC   expressed on the cell surface. The binding is preferably perform
CC   vivo. APO4 polypeptides/ active fragments are also useful for sc
CC   for agonists and antagonists by binding and observing the change
CC   activity. Effective pharmacological agents useful in diagnosis c
CC   treatment of disease are also identified using APO4 polypeptides
CC   fragments and APO4 signal transducer molecules that specifically
CC   with a cytoplasmic domain of APO4 and detecting a change in leve
CC   activity. The method is performed in vivo or in vitro. APO4 polyp
CC   are all useful as immunogens for preparing antibodies. APO4 is a
CC   useful for diagnosis/treatment of developmental or gestational
CC   abnormalities. APO8 was transfected to human breast carcinoma ce
CC   MCF-7, and induced apoptosis
```

```
CTGGCGCTGGCCTGCCTCGGCCTTCCTGCTGCCGGTCAGTTTGGGGAGCCGGGGC 248
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTGGCCGCTGGCCTGCCTCGGCCTTCCTGCTGCCGGTCAGTTTGGGGAGCCGGGGC 143

CTGTCCGCCCAGGAGCCTGCCCAGGAGAGCTGGTGGCAGAGGACCAGGAGACCC 308
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTGTCCGCCCAGGAGCCTGCCCAGGAGAGCTGTGGCAGAGGAGGACCAGGAGACCC 203

GAACTGAATCCCCAGACAGAGAAGACCAGGATCCTGCGCCTTTCCTGAACCGACT 368
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GAACTGAATCCCCAGACAGAGAAGCCAGGATCCTGCGCCTTTCCTGAACCGACT 263

CGGCCTCGCGAGAAGTGCACCTAAAGCCCGGAAAAACACCGGCTCGAAGAGCGATCGC 428
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CGGCCTCGAAGAAGTGCACCTAAAGCCCGGAAAAACACCGGGCTCGAAGAGCGATCGC 323

CATTATGAAGTTCATCCACGACCTGGACAGGACGGAGCGCAGGCAGGTGTGGACGG 488
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
CATTATGAAGTTCATCCACGACCTGGACAGGACGGAGCGCAGGCAGGTGTGGACGG 383

AGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTTCCAGCCCTCTGCGCTACAACCG 548
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTTCCAGCCCTCTGCGCTACAACCG 443

GATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCACTACCGTGTACTGTCAGGTGCA 608
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCACTACCGTGTACTGTCAGGTGCA 503

TGATGAGGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTGTGTGGATGGTGCTGGC 668
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGATGAGGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTGTGTGGATGGTGCTGGC 563

GCGGCTGCCTGGAGGAATTCTCAGCCACTGCGCGCCAGTTCCCTCGGGCCCCAGCTCCG 728
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCGGCTGCCTGGAGGAATTCTCAGCCACTGCGCGCCGAGTTCCCTCGGGCCCCAGCTCCG 623

CTGCCAGGTGTCTCGGGCGTGTTGGGCCCTGCGGCCAGGGTCTCCCTTGCGGATCCGCAC 788
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTGCCAGGTGTCTCGGGCGTGTTGGGCCCTGCGGCCAGGGTCTCCCTTGCGGATCCGCAC 683

CCCCTGGCCCATCTCAAGGGCTGCCCCCTTCCTCACCTACTTCGGACTCTTCCAGGT 848
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCCTGGCCCATCTCAAGGGCTGCCCCCTTCCTCACCTACTTCGGACTCTTCCAGGT 743

CTGAGGGGCCCGGCTCCCCCACAGTCGTCCCGGCGTGCCGCCGCTCCCCTCGACAGCT 908
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTGAGGGGCCCGGTCTCCCCGCAGTCGTCCCGGCGTGCCGCCGCTCCCCTCGACAGCT 803

TGGGCCACCCGGTCCCCTCCCCTCGCCCACCCTCAGCCGCGCTCTTTGCTCCAGACCTGCCCC 968
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGGGCCACCCGGTCCCCTCCCCTCGCCCACCCTCAGCCGCGCTCTTTGCTCCAGACCTGCCCC 863

CTCTAGAGGCTGCCTGGGCTTCAGTGTTTTCCATCCCACATAAATACAGTATT 1028
|||||||||||||||||||||||||||||||||||||||||||||||||||||
CTCTAGAGGCTGCCTGGGCCTGTTCACGTGTTTTCATCCCACATAAATACAGTATT 923

ACTCTTATCTTACAACTCCCCACCGCCCACTCCCACCTCCACTAGCTCCCCAATCC 1088
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
ACTCTTATCTTACAACAACACCCCACCGCCCACTCCCACCTCCACTAGCTCCCCAATCC 983

ACCCCTTGAGGCCCCAGTGATCTCGACTCCCCCTGGCCACGAACCCCCAGGGCAT 1148
||||||||||||||||||||||||||||||||||||||||||||||||||||||
ACCCCTTGAGGCCCCAGTGATCTCGACTCCCCCTGGCCACGAACCCCCAGGGCAT 1043

GTTCACTGTACTCTGTGGGCAAGGATGGTCCAGAAGACCCCACTTCAGGCACTAAG 1208
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTTCACTGTACTCTGTGGGCAAGGATGGTCCAGAAGACCCCACTTCAGGCACTAAG 1103

GGCTGGACCTGACCTGGCGGCAGGAAGCCAAAGACTGGGCCTAGGCCAGGAGTTCCCAAA 1268
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGCTGGACCTGGCGGCAGGAAGCCAAAGACTGGGCCTAGGCCAGGAGTTCCCCAAA 1163

GAGGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAA 1328
```

```
XX
KW   TWEAK extracellular domain; tumour necrosis factor; TNF; angiogen
KW   ocular neovascularisation; diabetic retinopathy; neovascular glau
KW   retinoblastoma; retinopathy of prematurity; retrolental fibroplas
KW   rheumatism; di
KW   rubeosis; uveitis; macular degeneration; arthritis;
KW   corneal graft neovascularisation; psoriasis; metastatic condition
KW   malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic
KW   preneoplastic condition; myocardial angiogenesis; wound granulati
KW   scleroderma; vascular adhesion; telangiectasia; ischaemia; human;
KW   atherosclerotic plaque neovascularisation; coronary atheroscleros
KW   peripheral atherosclerosis; pDC409-LZ-TWEAK; TWEAK receptor; TWEA
KW   fusion protein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             52. .873
FT                   /*tag= a
FT                   /product= "Fusion protein comprising a growth hor
FT                   leader, a leucine zipper multimerisation domain,
FT                   human TWEAK extracellular domain"
XX
PN   WO200145730-A2.
XX
PD   28-JUN-2001.
XX
PF   19-DEC-2000; 2000WO-US034755.
XX
PR   20-DEC-1999;  99US-0172878P.
PR   10-MAY-2000; 2000US-0203347P.
XX
PA   (IMMV ) IMMUNEX CORP.
XX
PI   Wiley SR;
XX
DR   WPI; 2001-417975/44.
DR   P-PSDB; AAU03499.
XX
PT   Modulating angiogenesis in a mammal for treating diseases mediate
PT   angiogenesis, e.g. solid tumors and vascular deficiencies of card
PT   peripheral tissue, by administering antagonist or agonist of TWEA
PT   receptor.
XX
PS   Example 1; Page 39-40; 46pp; English.
XX
CC   The sequence represents a DNA from the expression vector pDC409-L
CC   which encodes a fusion protein comprising a growth hormone leader
CC   leucine zipper multimerisation domain, and the extracellular doma
CC   human TWEAK. The fusion protein was used in the isolation of huma
CC   receptor (TWEAKR)-expressing clones from a COS cell human cDNA li
CC   The TWEAK protein is a member of the tumour necrosis factor (TNF)
CC   and induces angiogenesis. TWEAKR may therefore be used to screen
CC   develop TWEAK agonists and antagonists for the modulation of
CC   angiogenesis, to be used in the treatment and diagnosis of human
CC   The disorders mediated by angiogenesis include ocular disorders
CC   characterised by ocular neovascularisation such as diabetic retir
CC   neovascular glaucoma, retinoblastoma, retinopathy of prematurity,
CC   retrolental fibroplasia, rubeosis, uveitis, macular degeneration
CC   corneal graft neovascularisation, and inflammatory diseases such
CC   arthritis, rheumatism and psoriasis. Other treatable diseases inc
CC   malignant and metastatic conditions such as sarcomas and carcinom
CC   benign tumours and preneoplastic conditions, myocardial angiogene
CC   haemophilic joints, scleroderma, vascular adhesions, atherosclerc
CC   plaque neovascularisation, telangiectasia, wound granulation, cor
CC   atherosclerosis, peripheral atherosclerosis and ischaemia
XX
SQ   Sequence 898 BP; 187 A; 266 C; 267 G; 178 T; 0 U; 0 Other;
```

Query Match   45.5%; Score 625; DB 4; Length 898;
Best Local Similarity 100.0%; Pred. No. 6e-281;
Matches 625; Conservative 0; Mismatches 0; Indels 0; C

BP; 223 A; 317 C; 279 G; 211 T; 0 U; 0 Other;

   57.1%; Score 784; DB 2; Length 1030;
 rity 99.9%; Pred. No. 0;
 nservative 0; Mismatches 0; Indels 1; Gaps 0;

```
TTTGGGGAGCCGGCATGCGTCGTCCGCCCAGGAGCCTGCCCAGGAGGAGCTGGTG 288
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TTTGGGGAGCCGGCATGCGTCGTCCGCCCAGGAGCCTGCCCAGGAGGAGCTGGTG 60

GGGAGGACCAGGACCCGTCGGAACTGAATCCCCAGACAGAAGAAAGCCAGGATCCT 348
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGGAGGACCAGGACCCGTCGGAACTGAATCCCCAGACAGAAGAAAGCCAGGATCCT 120

TTTCCTGAACCGACTAGTTCGGCCTTCGCGAGAAGTGCACCTAAAGGCCGGAAAACA 408
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TTTCCTGAACCGACTAGTTCGGCCTTCGCGAGAAGTGCACCTAAAGGCCGGAAAACA 180

TCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACCGA 468
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACCGA 240

GGCAGGTGTGTGGACGGGACAGTGAGTGGGCTGGGAGGAAGCCAGAATCAACAGCTCC 528
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGCAGGTGTGTGGACGGGACAGTGAGTGGGCTGGGAGGAAGCCAGAATCAACAGCTCC 300

TCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTAC 588
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCTGCGCTACTACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTAC 360

TGTACTGTCAGGTGCACTTTGATGAGGGGGAAGGCTGTCTACCTGAAGCTGGACTTG 648
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGTACTGTCAGGTGCACTTTGATGAGGGGGAAGGCTGTCTACCTGAAGCTGGACTTG 420

TGGGATGGTGTGTCGGCCTGGCCTGGCCCTGCGCGCTGGCCTGCTGCCTGAGGAATTCTCAGCCACTGCGGCCAGT 708
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGGGATGGTGTGTCGGCCTGGCCTGGCCCTGCGCGCTGGCCTGCTGCCTGAGGAATTCTCAGCCACTGCGGCCAGT 480

TCGGGGCCCAGCTCCGCTCCCTGCCAGTGTCTCGGGCTGTTGGCCCTGGCCTGGCCCTGGCCAGGG 768
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCGGGGCCCAGCTCCGCTCCCTGCCAGTGTCTCGGGCTGTTGGCCCTGGCCTGGCCCTGGCCAGGG 540

CCCTGCGGGATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCGCCCCCTTCCTCACC 828
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCTGCGGGATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCGCCCCCTTCCTCACC 600

TCGGGACTCTTCCAGGTTCACTGAGGGCCCTGGGCCCTGGTCCCCACAGTGTCCCAGGCT 888
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCGGGACTCTTCCAGGTTCACTGAGGGCCCTGGGCCCTGGTCCCCACAGTGTCCCAGGCT 660

GCTCCCCTGCGACAGCTCTCTGGGCACCCGGTCCCCCTGCCCCCTCAGCCGCGT 948
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTCCCCCTGCGACAGCTCTCTGGGCACCCGGTCCCCCTGCCCCCTCAGCCGCGT 720

GCTCCAGACCTGCCCCCTCCCTAGAGGCTGCCTGGGCCTGTTCACGTGTTTTCCA 1008
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCTCCAGACCTGCCCCCTCCCTAGAGGCTGCCTGGGCCTGTTCACGTGTTTTCCA 780

ACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCCACT 1063
|||||||||||||||||||||||||||||||||||||||||||||||||||
ACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCCACT 835
```

dard; DNA; 898 BP.

(first entry)

ctor pDC409-LZ-TWEAK fusion protein-encoding DNA.

XX
PI　Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX
DR　WPI; 2003-615964/58.
XX
PT　New polynucleotide sequences obtained from various cDNA libraries
PT　as hybridization probes, as oligomers for PCR, for chromosome and
PT　mapping, in the recombinant production of protein, or in generati
PT　antisense DNA or RNA.
XX
PS　Claim 1; SEQ ID NO 21225; 44pp; English.
XX
CC　The invention relates to an isolated polynucleotide comprising an
CC　38043 cDNA sequences, appearing as ACH12789-ACH50831, whose seque
CC　determined by the technique of SBH (sequencing by hybridisation).
CC　included is a purified polypeptide comprising a sequence correspo
CC　a reading frame of the novel polynucleotide. The nucleic acid seq
CC　are useful in diagnostics as expressed sequence tags (EST) for
CC　identifying expressed genes or for physical mapping of the human
CC　in forensics, in assessing biodiversity, or in identifying muta
CC　responsible for genetic disorders and other traits. The nucleotid
CC　sequences are also useful as hybridisation probes, as oligomers f
CC　for chromosome and gene mapping, in the recombinant production of
CC　protein, or in generating antisense DNA or RNA. The purified poly
CC　is useful for generating antibodies specific for it. The present
CC　is one of the 38043 isolated cDNA/EST sequences. Note: The sequen
CC　for this patent did not form part of the printed specification, b
CC　obtained in electronic format directly from USPTO at
CC　seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ　Sequence 493 BP; 87 A; 180 C; 120 G; 101 T; 0 U; 5 Other;

Query Match　　29.6%;　Score 407;　DB 8;　Length 493;
Best Local Similarity　100.0%;　Pred. No. 2.6e-179;
Matches　407;　Conservative　0;　Mismatches　0;　Indels　0;

Qy　874　CAGTCGTCCCAGGCTGCGCGGCTCCCCTCGACAGCTCTCTGGGCACCGGTCCCCT(
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　 87　CAGTCGTCCCAGGCTGCGCGGCTCCCCTCGACAGCTCTCTGGGCACCGGTCCCCT(

Qy　934　CCACCCTCAGCCGGCTCTTTGCTCCAGACCTGCCCCTCCCTCAGAGGCTGCCTGCG(
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　147　CCACCCTCAGCCGGCTCTTTGCTCCAGACCTGCCCCTCCCTCAGAGGCTGCCTGGC

Qy　994　TTCACGTGTTTCCATCCCACATAAATACAGTATTCCCACTCTTATCTTACAACT(
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　207　TTCACGTGTTTCCATCCCACATAAATACAGTATTCCCACTCTTATCTTACAACT(

Qy　1054　ACCGCCACTCTCCACCTCACTAGCTCCCCAATCCTGACCCTTTGAGGCCCCCA(
　　　　　 ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　267　ACCGCCCACTCTCCACCTCACTAGCTCCCCAATCCTGACCCTTTGAGGCCCCCA(

Qy　1114　CTCGACTCCCCCTGGCCACAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCA(
　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　327　CTCGACTCCCCCTGGCCACAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCA(

Qy　1174　GGATGGGTCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGCA(
　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　387　GGATGGGTCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGCA(

Qy　1234　CCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGA　1280
　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||
Db　447　CCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGA　493

RESULT 14
ABK29540
ID　ABK29540 standard; cDNA; 195 BP.
XX
AC　ABK29540;
XX
DT　23-APR-2002　(first entry)
XX

Qy　TGGGGAGCCGGGCATCGTGTCTCGCCCAGGAGGCCTGCCCAGGAGGAGCTGGTGGCA　291
　　 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　TGGGGAGCCGGGCATCGTGTCTCGCCCAGGAGGCCTGCCCAGGAGGAGCTGGTGGCA　309

Qy　AGGACCAGGACCCGTCGGAACTGAATCCCCAGACACAGAAGAAAGCCAGGATCCTGCG　351
　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　AGGACCAGGACCCGTCGGAACTGAATCCCCAGACACAGAAGAAAGCCAGGATCCTGCG　369

Qy　TCCTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAACACGG　411
　　 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　TCCTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAACACGG　429

Qy　GAAGAGGCGATCGCAGCCCATTATGAAGTCATCCACGACCTGACAGGACGGAGCG　471
　　 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　GAAGAGGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACGGAGCG　489

Qy　CAGGTGTGGACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAGC　531
　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　CAGGTGTGGACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAGC　549

Qy　TGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGCGTGGGCTCTACTAC　591
　　 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　TGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGCGTGGGCTCTACTAC　609

Qy　ACTGTCAGGTGCACTTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTG　651
　　 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　ACTGTCAGGTGCACTTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTG　669

Qy　ATGGTGTGCTGGCCCTGCCTGGAGGAATTCTCAGCCACTGCGGCCAGTTCC　711
　　 |||||||||||||||||||||||||||||||||||||||||||||||||||
Db　ATGGTGTGCTGGCCCTGCCTGGAGGAATTCTCAGCCACTGCGGCCAGTTCC　729

Qy　TGGGCCCCAGCTCCGGCTCCGCCCTGCCCAGGTGTCTGTTGGCCCTGCGCCGGCCAGGGTCC　771
　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　TGGGCCCCAGCTCCGGCTCCGCCCTGCCAGGTGTCTGTTGGCCCTGCGCCGGCCAGGGTCC　789

Qy　TTGCGGATCCGACACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCCACCTAC　831
　　 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　TTGCGGATCCGACACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCCACCTAC　849

Qy　TGACTCTTCCAGGTTCACTGAG　856
　　 ||||||||||||||||||||||
Db　TGACTCTTCCAGGTTCACTGAG　874

dard; cDNA; 493 BP.

(first entry)

lial cell cDNA #2146.

equencing by hybridisation; SBH; expressed sequence tag; EST;
g; biodiversity; genetic disorder.

-A1.

2001US-00918995.

2001US-00918995.

NAC R T.
T I.
HE-CRAIN B.
SON M C.
S L W.

cinoma-specific cDNA #66.

denocarcinoma; colon cancer; tumour; gene; ss.

001WO-US018574.

000US-0210667P.
000US-0252614P.

CORP.

ing GE, Xu J, Secrist H;

052/13.

polynucleotide encoding a polypeptide comprising a portion
protein, for detection, diagnosis and therapy of human

133; 211pp; English.

relates to an isolated polynucleotide (I) encoding a
I) comprising a portion of a colon tumour protein. A new
e (III) that hybridises to (I) is useful for determining
of a cancer in a patient. (II) or antigen presenting cells
are useful for stimulating and/or expanding T cells
tumour protein, by contacting T cells with (I), (II) or
ting cells that express (I). (I), (II), or antigen
ls that express (II) are useful for treating colon cancer
y incubating CD4+ and/or CD8+ T cells isolated from a
(I), (II), or antigen presenting cells that express (II), so
roliferate, and administering to the patient an effective
proliferated T cells, thus inhibiting the development of a
patient. A new composition is useful for stimulating an
e in a patient. (I) or (II) is useful in vaccines and
compositions for prevention and treatment of colon cancer
agnosis and monitoring of the cancers. (I), (II) or an
st (II) is useful for detection, diagnosis and/ or therapy
l cancer. (I) is useful as a probe or primer for nucleic
tion, and in the design and preparation of ribozyme
inhibiting expression of (II) in tumour cells. ABK29475-
esent human colon adenocarcinoma-specific cDNA sequences of

BP; 49 A; 51 C; 58 G; 37 T; 0 U; 0 Other;

13.0%;   Score 179; DB 6; Length 195;
arity 100.0%;   Pred. No. 5.3e-73;
onservative 0; Mismatches 0; Indels 0; Gaps 0;

```
TGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAAGAGACCCCACTTCAGGCACT 1205
||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAAGAGACCCCACTTCAGGCACT 76

GGGGCTGGACCTGGCCGGCAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCC 1265
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGGGCTGGACCTGGCCGGCAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCC 136

GTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTT 1324
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTT 195
```

dard; cDNA to mRNA; 282 BP.

AC   AAT22190.
XX
DT   27-AUG-1996   (first entry)
XX
DE   Human gene signature HUMGS03761.
XX
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
XX
OS   Homo sapiens.
XX
PN   WO9514772-A1.
XX
PD   01-JUN-1995.
XX
PF   11-NOV-1994;   94WO-JP001916.
XX
PR   12-NOV-1993;   93JP-00355504.
XX
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
XX
PI   Matsubara K, Okubo K;
XX
DR   WPI; 1995-206931/27.
XX
PT   Single-stranded DNA for identifying gene signatures - isolated fr
PT   directed human cDNA library that reflects relative abundance of c
PT   mRNA in specific human tissues.
XX
PS   Claim 1; Page 1067; 2245pp; Japanese.
XX
CC   A single-stranded DNA (or its complementary strand or the corresp
CC   -stranded DNA) which comprises one of the 7837 "GS" sequences giv
CC   AAT19001-T26837 and which is able to hybridise to part of human g
CC   DNA, cDNA or mRNA is claimed. The GS (Gene Signature) sequences w
CC   obtained from 3'-directed cDNA libraries prepared from various hu
CC   tissues; synthesis of cDNA was initiated from the 3'-end of mRNA
CC   poly(T) as the sole primer. Since the 3'- untranslated sequence i
CC   to a particular mRNA species, almost all the 3'-oriented cDNAs hy
CC   with specific mRNAs. Each library is constructed so as to reflect
CC   accurately the relative abundance of different mRNAs in the parti
CC   tissue from which it was derived. The appearance frequency of a g
CC   in a cDNA library can be determined (esp. using primers and probe
CC   derived from the GS sequences) as a means of diagnosing abnormal
CC   function or for recognising different cell types
XX
SQ   Sequence 282 BP; 80 A; 62 C; 69 G; 66 T; 0 U; 5 Other;

Query Match   5.5%;   Score 76; DB 2; Length 282;
Best Local Similarity   100.0%;   Pred. No. 5.4e-25;
Matches   76;   Conservative   0;   Mismatches   0;   Indels   0;   G

```
Qy   1146 CATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCCCACTTCAGG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     36 CATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCCCACTTCAGG

Qy   1206 AAGAGGGGCTGGACCT 1221
           ||||||||||||||||
Db     96 AAGAGGGGCTGGACCT 111
```

RESULT 16
ADC97713
ID   ADC97713 standard; cDNA; 1239 BP.
XX
AC   ADC97713;
XX
DT   15-JAN-2004   (first entry)
XX
DE   Murine FL-TWEAK coding sequence.
XX

EAK, TNF relatedness and weak ability to induce cell death;
ecrosis Factor; TWEAK; fibrosis; cardiac disease;
; lung disease; kidney disease; skin disease;
le disease; adipose tissue disease;
nal tract disease; pancreatic disease;
organ disease; neural disease; cartilage disease;
connective tissue disease; cellular death; hepatotropic;
l; gastrointestinal; osteopathic; gene; ss.

Location/Qualifiers
1..750
/*tag= a
/product= "FL-TWEAK"

A2.

2003WO-US011350.
2002US-0371611P.

N INC.

kubowski A, Zheng T, Hahm K;

256/78.
712.

3Q ID NO 2; 120pp; English.

sequence is the coding sequence for murine transmembrane FL-
platedness and weak ability to induce cell death, where TNF
crosis Factor). TWEAK is a member of the TNF family. TWEAK
antagonists are useful for treating a TWEAK-related
.g. fibrosis; cardiac disease; liver disease; lung disease;
se; skin disease; skeletal muscle disease; adipose tissue
rointestinal tract disease; pancreatic disease; reproductive
a; neural disease; cartilage disease; bone disease;
issue disease; cellular death; and a pathological condition
expressing a TWEAK receptor.

9 BP; 249 A; 386 C; 331 G; 273 T; 0 U; 0 Other;

larity 100.0%;   Score 64;   DB 9;   Length 1239;
Conservative 0;   Pred. No. 2e-19;   Indels 0;   Gaps 0;
   4.7%;   Mismatches 0;

```
GCGCCCGTCGGAGCCAGAGGCCGGAGGGGGGCCCCGGGGGGGAGCCGGGCACCGCCCCTG 165
         ||||||||||||||||||||||||||||||||||||||||||||||
GCCGCCCGTCGGAGCCAGAGGCCGAGGGGGCCGAGGGGGGCCCGGGGGGGAGCCGCACCGCCCCTG 60
```

G 169

G 64

ndard; DNA; 60 BP.

(first entry)

d transcript detection oligonucleotide SEQ ID NO:31596.

KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB001903.
XX
PR   28-JUL-2000; 2000US-0221607P.
PR   02-MAY-2001; 2001US-0287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription
PT   genome, useful for detecting tissue-, pathology-, and development
PT   specific genes.
XX
PS   Example 1; SEQ ID NO 31596; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for det
CC   messenger RNAs that populate a (sub-)transcriptome, where the (su
CC   )transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprise
CC   oligonucleotides, each capable of hybridising selectively to a se
CC   messenger RNAs transcribed from a given transcription unit of the
CC   which encodes one or more messenger RNA splice variants. The
CC   oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitati
CC   quantitatively characterising the corresponding transcriptome, a
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mi
CC   libraries to detect transcripts of a sub-transcriptome under a pa
CC   biological or pathological state, and so allowing the detection
CC   - and pathology-specific genes such as those genes only expresse
CC   specific tissue under a specific pathological condition; to dete
CC   developmental specific genes; and to detect RNA transcripts and
CC   variants of a transcriptome of a patient suffering from a partic
CC   disorder. ABN27253 to ABN59589 represent oligonucleotide sequenc
CC   rats, humans and mice, which are used in the exemplification of
CC   present invention. N.B. The sequence data for this patent did no
CC   part of the printed specification, but was obtained in electroni
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 60 BP; 13 A; 16 C; 17 G; 14 T; 0 U; 0 Other;

Query Match   4.4%;   Score 60;   DB 6;   Length 60;
Best Local Similarity   100.0%;   Pred. No. 1.7e-17;
Matches   60;   Conservative   0;   Mismatches   0;   Indels   0;

```
Qy 1145 GCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCCCACTTCA
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 GCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCCCACTTCA
```

RESULT 18
ABN41049
ID   ABN41049 standard; DNA; 60 BP.
XX
AC   ABN41049;
XX
DT   15-JUL-2002 (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:137
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;

; transcriptome; oligonucleotide library; ss.

XX
OS   Homo sapiens.
XX
PN   WO20210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB001903.
XX
PR   28-JUL-2000; 2000US-0221607P.
PR   02-MAY-2001; 2001US-0287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription u:
PT   genome, useful for detecting tissue-, pathology-, and development:
PT   specific genes.
XX
PS   Example 1; SEQ ID NO 31341; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for det:
CC   messenger RNAs that populate a (sub-)transcriptome, where the (su:
CC   )transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   oligonucleotides, each capable of hybridising selectively to a se:
CC   messenger RNAs transcribed from a given transcription unit of the
CC   which encodes one or more messenger RNA splice variants. The
CC   oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitativ:
CC   quantitatively characterising the corresponding transcriptome, an:
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised min:
CC   libraries to detect transcripts of a sub-transcriptome under a pa:
CC   biological or pathological state, and so allowing the detection o:
CC   - and pathology-specific genes such as those genes only expressed
CC   specific tissue under a specific pathological condition; to detec:
CC   developmental specific genes; and to detect RNA transcripts and s:
CC   variants of a transcriptome of a patient suffering from a particu:
CC   disorder. ABN27253 to ABN59589 represent oligonucleotide sequence:
CC   rats, humans and mice. N.B. The sequence data for this patent did not
CC   present invention. N.B. The sequence data for this patent did not
CC   part of the printed specification, but was obtained in electronic
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 60 BP; 12 A; 19 C; 10 G; 19 T; 0 U; 0 Other;

Query Match   4.4%;   Score 60; DB 6; Length 60;
Best Local Similarity   100.0%; Pred. No. 1.7e-17;
Matches 60; Conservative 0; Mismatches 0; Indels 0; G.

Qy   979 GGCTGCCTGGGCCTGTTCACGTGTTTCATCCCACATAAATACAGTATTCCCACT
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GGCTGCCTGGGCCTGTTCACGTGTTTCATCCCACATAAATACAGTATTCCCACT

RESULT 20
ABN58591
ID   ABN58591 standard; DNA; 60 BP.
XX
AC   ABN58591;
XX
DT   15-JUL-2002 (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:3133
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX

:001WO-IB001903.

:00US-0221607P.
:01US-0287724P.

:EN INC.

:sserman A,   Mintz E,   Mintz L,   Faigler S;

:83/30.

:otide libraries comprising oligonucleotides which
:bridize to mRNAs transcribed from a transcription unit of a
:for detecting tissue-, pathology-, and developmental-

: ID NO 13797; 47pp; English.

:vention describes oligonucleotide libraries for detecting
: that populate a (sub-)transcriptome, where the (sub-
: comprises messenger RNAs transcribed from multiple
: units that populate a genome. The library comprises several
:es, each capable of hybridising selectively to a set of
: transcribed from a given transcription unit of the genome,
: one or more messenger RNA splice variants. The
:le libraries are useful for detecting mRNAs from a
:ple, in expression profiling studies, in qualitatively or
:characterising the corresponding transcriptome, and in
:transcripts and splice variants of human or animal
:. The libraries may also be used as specialised mini
:letect transcripts of a sub-transcriptome under a particular
:pathological state, and so allowing the detection of tissue
:y-specific genes such as those genes only expressed in
:e under a specific pathological condition; to detect
:specific genes; and to detect RNA transcripts and splice
:transcriptome of a patient suffering from a particular
:7253 to ABN59589 represent oligonucleotide sequences from
:nd mice, which are used in the exemplification of the
:ion. N.B. The sequence data for this patent did not form
:inted specification, but was obtained in electronic format
:WIPO at ftp.wipo.int/pub/published_pct_sequences

'; 17 A; 15 C; 15 G; 13 T; 0 U; 0 Other;

rity   4.4%;   Score 60; DB 6; Length 60;
:   100.0%; Pred. No. 1.7e-17;
nservative 0; Mismatches 0; Indels 0; Gaps 0;

:ATGTGAGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTT 1322
:||||||||||||||||||||||||||||||||||||||||||||||||||||||
:ATGTGAGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTT 60

:lard; DNA; 60 BP.

:(first entry)

:transcript detection oligonucleotide SEQ ID NO:31341.

:rat; splice transcript; detection; RNA transcript;
:; transcriptome; oligonucleotide library; ss.

2.

XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF
PR 20-JUL-2001; 2001WO-IB001903.
XX
PR 28-JUL-2000; 2000US-0221607P.
PR 02-MAY-2001; 2001US-0287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription u
PT genome, useful for detecting tissue-, pathology-, and development
PT specific genes.
XX
PS Example 1; SEQ ID NO 31597; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for det
CC messenger RNAs that populate a (sub-)transcriptome, where the (su
CC)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC oligonucleotides, each capable of hybridising selectively to a se
CC messenger RNAs transcribed from a given transcription unit of the
CC which encodes one or more messenger RNA splice variants. The
CC oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitati
CC quantitatively characterising the corresponding transcriptome, ar
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mir
CC libraries to detect transcripts of a sub-transcriptome under a pr
CC biological or pathological state, and so allowing the detection o
CC - and pathology-specific genes such as those genes only expresse
CC specific tissue under a specific pathological condition; to detec
CC developmental specific genes; and to detect RNA transcripts and t
CC variants of a transcriptome of a patient suffering from a partic
CC disorder. ABN27253 to ABN59589 represent oligonucleotide sequenc
CC rats, humans and mice, which are used in the exemplification of t
CC present invention. N.B. The sequence data for this patent did not
CC part of the printed specification, but was obtained in electroni
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 60 BP; 17 A; 15 C; 15 G; 13 T; 0 U; 0 Other;

Query Match 4.4%; Score 60; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.7e-17;
Matches 60; Conservative 0; Mismatches 0; Indels 0;

Qy 1263 CCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTG
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTG

RESULT 22
AAX56002
ID AAX56002 standard; DNA; 50 BP.
XX
AC AAX56002;
XX
DT 15-JUL-1999 (first entry)
XX
DE Human tumour necrosis factor Apo-3 ligand probe SEQ ID NO:4.
XX
KW Human; tumour necrosis factor; Apo-3 ligand; lymphotoxin; apopto
KW NF-kappaB-dependent transcription; JNK/SAPK-dependent response;
KW hybridisation; probe; ss.
XX
OS Synthetic.

---

asserman A, Mintz E, Mintz L, Faigler S;

383/30.

eotide libraries comprising oligonucleotides which
hybridize to mRNAs transcribed from a transcription unit of a
l for detecting tissue-, pathology-, and developmental-
s.

Q ID NO 31339; 47pp; English.

invention describes oligonucleotide libraries for detecting
us that populate a (sub-)transcriptome, where the (sub-
e comprises messenger RNAs transcribed from multiple
units that populate a genome. The library comprises several
des, each capable of hybridising selectively to a set of
s transcribed from a given transcription unit of the genome,
one or more messenger RNA splice variants. The
de libraries are useful for detecting mRNAs from a
mple, in expression profiling studies, in qualitatively or
y characterising the corresponding transcriptome, and in
transcripts and splice variants of human or animal
s. The libraries may also be used as specialised mini
detect transcripts of a sub-transcriptome under a particular
gy-specific genes such as those genes only expressed in
sue under a specific pathological condition; to detect
specific genes; and to detect RNA transcripts and splice
transcriptome of a patient suffering from a particular
27253 to ABN59589 represent oligonucleotide sequences from
and mice, which are used in the exemplification of the
tion. N.B. The sequence data for this patent did not form
rinted specification, but was obtained in electronic format
n WIPO at ftp.wipo.int/pub/published_pct_sequences

BP; 17 A; 15 C; 15 G; 13 T; 0 U; 0 Other;

4.4%; Score 60; DB 6; Length 60;
larity 100.0%; Pred. No. 1.7e-17;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

AAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGGATTT 1322
|||||||||||||||||||||||||||||||||||||||||||||||||||||
AAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGGATTT 60

ndard; DNA; 60 BP.

(first entry)

d transcript detection oligonucleotide SEQ ID NO:31597.

; rat; splice transcript; detection; RNA transcript;
nt; transcriptome; oligonucleotide library; ss.

.

PA   (GETH) GENENTECH INC.

XX

PI   Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA;
PI   Napier MA, Pitti RM, Wood WI;

XX

DR   WPI; 2000-442668/38.

XX

PT   Novel composition to inhibit neoplastic cell growth or for treati
PT   in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,
PT   PRO224, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or PRO866.

XX

PS   Example 3; Page 98; 172pp; English.

XX

CC   The present sequence is that of a DNA probe based on an isolated
CC   expressed sequence tag showing homology to human Apo-2 ligand. Th
CC   was used to screen a human foetal kidney cDNA library to identify
CC   clone DNA30879-1152 (see AAA49717), which encodes human antitumou
CC   protein PRO207 (see AAY95338). A claimed method for inhibiting th
CC   of a tumour cell comprises exposing the tumor cell to PRO179, PRO
CC   PRO320, PRO219, PRO221, PRO224, PRO328, PRO301, PRO526, PRO362, P
CC   PRO509 or PRO866 (see AAY95337-49). The tumour is especially brea
CC   ovarian, renal, colorectal, uterine, prostate, lung, bladder or c
CC   nervous system cancer, melanoma or leukaemia. Nucleic acids encod
CC   PRO179 etc. are used in the recombinant production of antitumour
CC   polypeptides

XX

SQ   Sequence 50 BP; 10 A; 18 C; 13 G; 9 T; 0 U; 0 Other;

Query Match   3.6%;   Score 50;   DB 3;   Length 50;
Best Local Similarity   100.0%;   Pred. No. 7.7e-13;
Matches   50;   Conservative   0;   Mismatches   0;   Indels   0;   G

Qy   527   CCAGCCCTCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGG 576
               |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1   CCAGCCCTCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGG 50

RESULT 24

ABK40292
ID   ABK40292 standard; DNA; 50 BP.

XX

AC   ABK40292;

XX

DT   15-JUL-2002 (first entry)

XX

DE   Oligonucleotide probe for human PRO207 DNA.

XX

KW   Human; PRO; benign tumour; malignant tumour; lymphoid malignancy;
KW   leukaemia; neuronal disorder; stromal disorder; blastocoelic disc
KW   inflammatory disorder; immune disorder; angiogenic disorder; cyto
KW   neuroprotective; probe; ss.

XX

OS   Homo sapiens.

XX

PN   WO20015348G-A1.

XX

PD   26-JUL-2001.

XX

PF   11-FEB-2000; 2000WO-US003565.

XX

PR   08-MAR-1999;   99WO-US005028.
PR   11-MAR-1999;   99US-0123972P.
PR   11-MAY-1999;   99US-0133459P.
PR   02-JUN-1999;   99WO-US012252.
PR   22-JUN-1999;   99US-0140650P.
PR   22-JUN-1999;   99US-0140653P.
PR   20-JUL-1999;   99US-0144758P.
PR   26-JUL-1999;   99US-0145698P.
PR   28-JUL-1999;   99US-0146222P.
PR   17-AUG-1999;   99US-0149395P.
PR   31-AUG-1999;   99US-0151689P.
PR   01-SEP-1999;   99WO-US020111.
PR   15-SEP-1999;   99WO-US021090.

(second column, partial text at left of page)

98WO-US021407.

97US-0062037P.
97US-0069862P.

TECH INC.

Marsters SA, Pitti R;

982/24.

3- ligand (a tumor necrosis factor) homologue.

ge 36; 74pp; English.

invention describes a human tumour necrosis factor (TNF) and
omologue designated Apo-3 ligand. Apo-3 ligand has
civity. Apo-3 ligand can be used to induce apoptosis in
cer cells, to induce NF-kappaB-dependent transcription and
/SAPK-dependent responses in mammalian cells. The present
esents an Apo-3 ligand probe, which is used in an example
ent invention

P; 10 A; 18 C; 13 G; 9 T; 0 U; 0 Other;

arity   100.0%;   Pred. No. 7.7e-13;
onservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

CCCTCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGG 576
||||||||||||||||||||||||||||||||||||||||||||||
CCCTCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGG 50

dard; DNA; 50 BP.

(first entry)

DNA probe.

; antitumour; tumour; therapy; cytostatic; breast cancer;
r; renal cancer; colorectal cancer; uterine cancer;
er; lung cancer; bladder cancer;
us system cancer; melanoma; leukaemia; neoplasm; probe; ss.

2.

99WO-US028565.

98US-0113296P.
99WO-US005028.
99US-0130232P.
99US-0131445P.
99US-0134287P.
99US-0144758P.
99US-0145698P.
99WO-US021090.
99WO-US021547.

XX
PT New system for leukocyte expression profiling, diagnosing a disea
PT monitoring (the rate of) progression of a disease, e.g. atherosc
PT or congestive heart failure, comprises diagnostic oligonucleotide
XX
PS Claim 1; Page 408; 0pp; English.
XX
CC The invention relates to a system for detecting gene expression,
CC comprises one or two isolated DNA molecules that detect expressio
CC gene, where the gene corresponds to any of 8143 oligonucleotides
CC (ABZ00010-ABZ08152) each having 50 base pairs (bp). The system i
CC for leukocyte expression profiling. It is particularly useful for
CC diagnosing a disease, monitoring (rate of) progression of a disea
CC predicting therapeutic outcome, determining prognosis for a pati
CC predicting disease complications in an individual or monitoring i
CC to treatment in an individual. The diseases include cardiac allo
CC rejection, kidney allograft rejection, liver allograft rejection,
CC atherosclerosis, congestive heart failure, systemic lupus erythe
CC rheumatoid arthritis, osteoarthritis or cytomegalovirus infectio
XX
SQ Sequence 50 BP; 14 A; 11 C; 19 G; 6 T; 0 U; 0 Other;

Query Match   3.6%;   Score 50;   DB 6;   Length 50;
Best Local Similarity   100.0%;   Pred. No. 7.7e-13;
Matches   50; Conservative   0; Mismatches   0; Indels   0; ○

Qy   1196 TTCAGGCACTAAGAGGGGCTGGACCTGGCGGCAGGAAGCCAAAGAGACTG 1245
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TTCAGGCACTAAGAGGGGCTGGACCTGGCGGCAGGAAGCCAAAGAGACTG 50

RESULT 26
AAX23425
ID AAX23425 standard; DNA; 701 BP.
XX
AC AAX23425;
XX
DT 18-JUN-1999 (first entry)
XX
DE Mouse TNRL3 DNA.
XX
KW Tumour necrosis factor receptor; signal transducer molecule; TNF
KW developmental abnormality; gestational abnormality; prostate c
KW APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy;
KW cytoplasmic domain; immunogen; antibody preparation; breast carc
KW apoptosis; mouse; ss.
XX
OS Mus sp.
XX
FH Key   Location/Qualifiers
FT CDS   1..636
FT   /*tag= a
FT   /product= "TNRL3"
XX
PN WO9911791-A2.
XX
PD 11-MAR-1999.
XX
PF 04-SEP-1998; 98WO-US018393.
XX
PR 05-SEP-1997; 97US-00924634.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Chaudhary PM;
XX
DR WPI; 1999-205191/17.
DR P-PSDB; AAW93591.
XX
PT New Tumor Necrosis Factor family receptor polypeptides and ligan
PT useful for diagnosis and treatment of prostate cancer and develo
PT or gestational abnormalities.
XX

PR 99WO-US028313.
PR 99WO-US028301.
PR 99WO-US028634.
PR 2000WO-US000219.
XX
PA TECH INC.
XX
PI Goddard A, Godowski PJ, Gurney AL, Hillan KJ,
PI Pan J, Pitti RM, Roy MA, Smith V, Stone DM;
PI Wood WI;
XX
...567/26.
XX
...nucleic acids encoding PRO polypeptides, useful for treating
...ignant tumors, leukemias and lymphoid malignancies,
...angiogenic and immunologic disorders.
XX
...ge 109; 302pp; English.
XX
CC ...nvention relates to the isolation of novel human PRO
CC (AAU86128-AAU86162) and the polynucleotide sequences
CC ... The PRO polypeptides, agonists, antagonists or anti-PRO
CC ...e useful for treating benign or malignant tumours (e.g.
CC ..., bladder, breast, etc), leukaemias and lymphoid
CC ...other disorders such as neuronal, glial, astrocytal,
CC ...glandular, macrophagal, stromal and blastocoelic disorders,
CC ...immune and angiogenic disorders. The polynucleotide
CC ...a also useful in gene therapy. The present sequence
CC ...probe used in the methods of the present invention
XX
SQ ...BP; 10 A; 18 C; 13 G; 9 T; 0 U; 0 Other;

...arity   100.0%;   Score 50;   DB 6;   Length 50;
...onservative   0; Mismatches   0; Indels   0; Gaps   0;

          3.6%;   Pred. No. 7.7e-13;

...CCCTCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGG 576
          ||||||||||||||||||||||||||||||||||||||||||||||
...CCCTCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGG 50

...idard; DNA; 50 BP.

...(first entry)

...yte gene expression profiling probe SEQ ID NO 2558.

...e; gene expression profiling; allograft rejection;
...sis; congestive heart failure; systemic lupus erythematosus;
...cthritis; osteoarthritis; cytomegalovirus; infection; probe;

A2.

2001WO-US047856.

2000US-0241994P.
2001US-0296764P.

ARDIA INC.

..., Fry K, Matcuk G, Altman P, Prentice J, Phillips J;
...ard R, Quertermous T, Johnson F;

6525/68.

Fig 13B; 156pp; English.

n describes isolated Tumor Necrosis Factor (TNF) family
peptides: APO4, APO6, APO8 and APO9 or their active
d isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
fragments. APO4 is useful for diagnosing prostate cancer by
evels of APO4 in an individual. Prostate cancer can also be
APO4 selective binding agents linked to a therapeutic
polypeptides are also useful for identifying selective
s, useful in diagnosis/treatment of disease by binding of
polypeptide/active fragment which is extracellular, or
the cell surface. The binding is preferably performed in
lypeptides/ active fragments are also useful for screening
and antagonists by binding and observing the changer in APO4
ective pharmacological agents useful in diagnosis or
disease are also identified using APO4 polypeptides/active
asmic signal transducer molecules that specifically interact
. APO4 signal domain of APO4 and detecting a change in level of APO4
method is performed in vivo or in vitro. APO polypeptides
l as immunogens for preparing antibodies. APO4 is also
agnosis/treatment of developmental or gestational
. APO8 was transfected to human breast carcinoma cell line
duced apoptosis

BP; 139 A; 210 C; 203 G; 149 T; 0 U; 0 Other;

arity 3.4%; Score 46; DB 2; Length 701;
onservative 100.0%; Pred. No. 5e-11; Indels 0; Gaps 0;
0; Mismatches 0;

```
TGGGCTCTACTACCTGTACTGTCAGGTGCACTTTGATGAGGG 620
||||||||||||||||||||||||||||||||||||||||||
TGGGCTCTACTACCTGTACTGTCAGGTGCACTTTGATGAGGG 401
```

dard; cDNA; 1168 BP.

(first entry)

tumour necrosis factor related ligand (TRELL) gene.

necrosis factor related ligand; tnf; treatment; cancer;
sease; immune system; stimulation; suppression;
on; db.

Location/Qualifiers
2. .679
/*tag= a
/note= "tumour necrosis factor related ligand"

97WO-US013945.

96US-0023541P.
96US-0028515P.
97US-0040820P.

N INC.
GENEVA FACULTY MEDICINE.

e Y, Browning JL;

619/13.
524.

XX
PT Tumour necrosis factor related ligand - useful for, e.g. treating
PT auto-immune disease and immune responses to tissue grafts.
XX
PS Claim 2; Page 45-46; 69pp; English.
XX
CC The sequence is that encoding mouse tumour necrosis factor relate
CC (TRELL). TRELL or active fragments can be included with a carrier
CC pharmaceutical compositions to treat cancer, autoimmune diseases
CC immune responses to tissue grafts, or to stimulate or suppress tl
CC system. It is useful to screen for TRELL receptors, by labelling
CC detectable label and screening compositions for binding. Agents
CC interfering with TRELL-receptor binding can also be screened for,
CC then be administered, optionally with interferon- gamma, to induc
CC death or treat, suppress or alter immune responses (especially in
CC human adenocarcinoma cells) involving a signal pathway between TP
CC its receptor. The DNA sequence can be used in gene therapy for TP
CC related disorders in mammals (especially humans), e.g. tumours,
CC autoimmune and inflammatory diseases or inherited genetic disorde
CC introducing into cells, and expressing, therapeutically effective
CC of a vector, e.g. a virus comprising a gene encoding TRELL. It ma
CC be of use in the preparation of prepare probes for screening
CC natural/synthetic DNAs for TRELL-encoding sequences and for anti
CC therapy
XX
SQ Sequence 1168 BP; 242 A; 360 C; 298 G; 268 T; 0 U; 0 Other;

Query Match   3.4%; Score 46; DB 2; Length 1168;
Best Local Similarity 100.0%; Pred. No. 4.9e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; C

```
Qy  575 GGGCTGGGCTCTACTACCTGTACTACCTGTACTGTCAGGTGCACTTTGATGAGGG 620
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  399 GGGCTGGGCTCTACTACCTGTACTACCTGTACTGTCAGGTGCACTTTGATGAGGG 444
```

RESULT 28
ABX37032
ID ABX37032 standard; cDNA; 408 BP.
XX
AC ABX37032;
XX
DT 20-FEB-2003 (first entry)
XX
DE Bovine EST associated with lactation/muscle/fat deposition #2197.
XX
KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identific
KW gene analysis; cattle breeding.
XX
OS Bos Taurus.
XX
PN US2002137139-A1.
XX
PD 26-SEP-2002.
XX
PF 24-SEP-2001; 2001US-00960352.
XX
PR 12-JAN-1999; 99US-0115707P.
PR 11-JAN-2000; 2000US-00480902.
XX
PA (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
XX
PI Byatt JC, Mathialagan N, Tao N, Warren WC;
XX
DR WPI; 2003-110599/10.
XX
PT New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and ar
PT cattle breeding, or for genetically improving cattle.

ID NO 2197; 245pp; English.

n relates to a purified nucleic acid molecule associated with
muscle and fat deposition (designated LMFD), derived from
the LMFD nucleic acid can specifically hybridise to a second
molecule comprising any of 15112 nucleotide sequences,
ABX34836-ABX49947, or complements of them. Also included are
sformed cell having a nucleic acid comprising an LMFD nucleic
to a promoter and a 3' non- translated sequence that
the cell to cause termination of transcription and addition
lated ribonucleotides to a 3' end of the mRNA molecule; and
ing a level or pattern of a molecule in a bovine cell or
ising: (a) incubating a marker nucleic acid (comprising any
y nucleic acid sequences or its complement or fragment) with a
y nucleic acid molecule obtained from the bovine cell or
e hybridisation between the marker nucleic acid and the
y nucleic acid permits the detection of the molecule; and (b)
e level or pattern of the complementary nucleic acid, where
n of the complementary nucleic acid is predictive of the
tern of the molecule. The LMFD nucleic acid is used for
a level or pattern of a molecule in a bovine cell or tissue.
for genome mapping, gene identification and analysis, cattle
eparation of constructs for use in cattle gene expression, or
ly improving cattle. The present sequence is one of the
LMFD EST (expressed sequence tag) nucleic acids. Note: The
ence was not shown in the specification but was obtained in
ormat from the USPTO web site:
o.gov/sequence.html?DocID=20020137139

BP; 78 A; 136 C; 137 G; 57 T; 0 U; 0 Other;

larity   100.0%;   Pred. No. 2.3e-06;   Length 408;
Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
GGCCTGCCTCGGCCTCCTCCTGCTGGCCGTGGTCAG 233
||||||||||||||||||||||||||||||||||||
GGCCTGCCTCGGCCTCCTCCTGCTGGCCGTGGTCAG 152
```

ndard; DNA; 412 BP.

(first entry)

ated gene, SEQ ID 1352.

; gene expression profile; hepatotoxicity; liver;
ng; toxicity assay; ds.

; LOGIC INC.

Porter M, Johnson K, Higgs B, Castle A, Elashoff M;

-A2.

2003WO-US003194.

2002US-00060087.
2002US-0364045P.
2002US-0364055P.
2002US-0436643P.

39530/65.

PT Predicting a toxic effect of a compound, useful in identifying t
PT markers in liver tissues or cells for drug screening and toxicit
PT comprises preparing gene expression profile of tissue or cells e
PT the compound.
XX
PS Claim 1; SEQ ID NO 1352; 1156pp; English.
XX
CC The present invention relates to a method for predicting a toxic
CC of a compound. The method comprises preparing a gene expression
CC of a tissue or cell sample exposed to the compound, and comparin
CC gene expression profile to a database comprising SEQ ID 1-4925,
CC differential expression of the gene indicates at least one toxic
CC The method is useful for predicting at least one toxic effect of
CC compound, predicting hepatotoxicity or the progression of a toxi
CC of a compound, identifying an agent that modulates the onset or
CC progression of a toxic response, predicting the cellular pathway
CC compound modulates in a cell, and identifying an agent that modu
CC least one activity of a protein. The method and compositions of
CC present invention using a database of genes having liver toxin-i
CC differential expression, are useful in identifying toxicity mark
CC liver tissues or cells for drug screening and toxicity assays. N
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly fr
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 412 BP; 104 A; 87 C; 116 G; 105 T; 0 U; 0 Other;

Query Match   2.1%; Score 29; DB 9; Length 412;
Best Local Similarity   100.0%; Pred. No. 0.0043;
Matches   29; Conservative   0; Mismatches   0; Indels   0;

```
Qy    1336 TATTTTTATTATTATTGTGACAAAATGTT 1364
           |||||||||||||||||||||||||||||
Db      67 TATTTTTATTATTATTGTGACAAAATGTT 39
```

RESULT 30
ABK40356/c
ID   ABK40356 standard; DNA; 26 BP.
XX
AC   ABK40356;
XX
DT   15-JUL-2002 (first entry)
XX
DE   Reverse PCR primer for gene amplification analysis of human PRO2
XX
KW   Human; PRO; benign tumour; malignant tumour; lymphoid malignancy
KW   leukaemia; neuronal disorder; stromal disorder; blastocoelic dis
KW   inflammatory disorder; immune disorder; angiogenic disorder; cyt
KW   neuroprotective; PCR; primer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153486-A1.
XX
PD   26-JUL-2001.
XX
PF   11-FEB-2000; 2000WO-US003565.
XX
PR   08-MAR-1999; 99WO-US005028.
PR   11-MAR-1999; 99US-0123972P.
PR   11-MAY-1999; 99US-0133459P.
PR   02-JUN-1999; 99WO-US012252.
PR   22-JUN-1999; 99US-0140650P.
PR   22-JUN-1999; 99US-0140653P.
PR   20-JUL-1999; 99US-0144758P.
PR   26-JUL-1999; 99US-0145698P.
PR   28-JUL-1999; 99US-0146222P.
PR   17-AUG-1999; 99US-0149395P.
PR   31-AUG-1999; 99US-0151689P.
PR   01-SEP-1999; 99WO-US020111.
PR   15-SEP-1999; 99WO-US021090.
PR   30-NOV-1999; 99WO-US028313.

99WO-US028301.
99WO-US028634.
2000WO-US000219.

TECH INC.

Goddard A, Godowski PJ, Gurney AL, Hillan KJ;
Pan J, Pitti RM, Roy MA, Smith V, Stone DM;
Wood WI;

567/26.

ucleic acids encoding PRO polypeptides, useful for treating
ignant tumors, leukemias and lymphoid malignancies,
angiogenic and immunologic disorders.

age 140; 302pp; English.

nvention relates to the isolation of novel human PRO
(AAU86128-AAU86162) and the polynucleotide sequences
. The PRO polypeptides, agonists, antagonists or anti-PRO
e useful for treating benign or malignant tumours (e.g.
, bladder, breast, etc), leukaemias and lymphoid
other disorders such as neuronal, glial, astrocytal,
glandular, macrophagal, stromal and blastocoelic disorders,
immune and angiogenic disorders. The polynucleotide
also useful in gene therapy. The present sequence
PCR primer used in the methods of the present invention

P; 7 A; 1 C; 9 G; 9 T; 0 U; 0 Other;

arity   100.0%; Pred. No. 0.12;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

TCCCACATAAATACAGTATTCC 1030
||||||||||||||||||||||
TCCCACATAAATACAGTATTCC 1

dard; DNA; 140 BP.

(first entry)

ssion-related sequence, SEQ ID 542.

irucide; apoptotic; gene therapy; tumour suppression;
ion; apoptosis; virus resistance; viral infection; tumour;
tive disease; ds.

2002WO-FR000543.

2001FR-00001925.

ULAR ENGINES LAB.

Amson R, Tuijnder M, Susini L;

286/05.

cid encoding a translationally controlled tumor protein,
eating, preventing and diagnosing viral, tumor or
diseases.

XX
PS   Disclosure; Page; 45pp; French.
XX
CC   The present invention relates to novel nucleic acid sequences (AF
CC   ABZ79313), which are involved in the molecular pathways of tumour
CC   suppression, tumour reversion, apoptosis and/or virus resistance.
CC   sequences are also useful for treatment or prevention of viral, t
CC   and cell degenerative diseases, and also for diagnosis and progno
CC   these diseases. Note: The sequence data for this patent is not
CC   represented in the printed specification but is based on sequence
CC   information supplied by the European Patent Office
XX
SQ   Sequence 140 BP; 43 A; 23 C; 28 G; 46 T; 0 U; 0 Other;

Query Match   1.9%; Score 26; DB 7; Length 140;
Best Local Similarity   100.0%; Pred. No. 0.11;
Matches   26; Conservative   0; Mismatches   0; Indels   0; G

Qy   1 ATGTCATTGTTAGACTTGAAATTC 26
        ||||||||||||||||||||||||
Db   67 ATGTCATTGTTAGACTTGAAATTC 42

RESULT 32
ABZ09382/c
ID   ABZ09382 standard; DNA; 140 BP.
XX
AC   ABZ09382;
XX
DT   16-JAN-2003   (first entry)
XX
DB   Human oligonucleotide SEQ ID 542.
XX
KW   Human; tumour suppressor; virucide; cytostatic; nootropic;
KW   neuroprotective; neuroleptic; gene therapy; tumour suppression;
KW   tumour reversion; apoptosis; viral resistance; viral infection;
KW   cell degeneration; Alzheimer's disease; schizophrenia; cancer; ds
XX
OS   Homo sapiens.
XX
PN   FR2822475-A1.
XX
PD   27-SEP-2002.
XX
PF   20-MAR-2002; 2002FR-00003459.
XX
PR   13-FEB-2001; 2001FR-00001925.
XX
PA   (MOLE-) MOLECULAR ENGINES LAB SA.
XX
PI   Telerman A, Amson R, Tuijnder M, Susini L;
XX
DR   WPI; 2003-032204/03.
XX
PT   New human nucleic acid, useful for diagnosis, prognosis and treat
PT   e.g. of tumors, also related vectors, transformed cell, polypepti
PT   antibodies.
XX
PS   Disclosure; Page 120; 189pp; French.
XX
CC   The present invention relates to human oligonucleotides (ABZ08841
CC   ABZ09860). The expression of the oligonucleotide is implicated i
CC   suppression or reversion, apoptosis and/or viral resistance. The
CC   oligonucleotides are useful for preventing and/or treating viral
CC   infection, tumour development and cell degeneration (e.g. Alzheim
CC   disease and schizophrenia), especially cancer
XX
SQ   Sequence 140 BP; 43 A; 23 C; 28 G; 46 T; 0 U; 0 Other;

Query Match   1.9%; Score 26; DB 7; Length 140;
Best Local Similarity   100.0%; Pred. No. 0.11;
Matches   26; Conservative   0; Mismatches   0; Indels   0; G

```
KW  cell degenerative disease; ds.
XX
OS  Unidentified.
XX
PN  WO200264731-A2.
XX
PD  22-AUG-2002.
XX
PF  13-FEB-2002; 2002WO-FR000543.
XX
PR  13-FEB-2001; 2001FR-00001925.
XX
PA  (MOLE-) MOLECULAR ENGINES LAB.
XX
PI  Telerman A, Amson R, Tuijnder M, Susini L;
XX
DR  WPI; 2003-058286/05.
XX
PT  New nucleic acid encoding a translationally controlled tumor pro
PT  useful for treating, preventing and diagnosing viral, tumor or
PT  degenerative diseases.
XX
PS  Disclosure; Page; 45pp; French.
XX
CC  The present invention relates to novel nucleic acid sequences (A
CC  ABZ79313), which are involved in the molecular pathways of tumou
CC  suppression, tumour reversion, apoptosis and/or virus resistance
CC  sequences are also useful for treatment or prevention of viral,
CC  and cell degenerative diseases, and also for diagnosis and progn
CC  these diseases. Note: The sequence data for this patent is not
CC  represented in the printed specification but is based on sequenc
CC  information supplied by the European Patent Office
XX
SQ  Sequence 145 BP; 51 A; 28 C; 23 G; 43 T; 0 U; 0 Other;

Query Match      1.9%;     Score 26;   DB 7;    Length 145;
Best Local Similarity   100.0%;   Pred.No. 0.11;
Matches    26;   Conservative   0;   Mismatches   0;   Indels   0;

Qy     1 ATGTCATTGTTAGACTTTGAAATTTC 26
         ||||||||||||||||||||||||||
Db    74 ATGTCATTGTTAGACTTTGAAATTTC 99

RESULT 35
ABZ08855
ID  AB208855  standard; DNA; 145 BP.
XX
AC  ABZ08855;
XX
DT  16-JAN-2003  (first entry)
XX
DE  Human oligonucleotide SEQ ID 15.
XX
KW  Human; tumour suppressor; virucide; cytostatic; nootropic;
KW  neuroprotective; neuroleptic; gene therapy; tumour suppression;
KW  tumour reversion; apoptosis; viral resistance; viral infection;
KW  cell degeneration; Alzheimer's disease; schizophrenia; cancer; d
XX
OS  Homo sapiens.
XX
PN  FR2822475-A1.
XX
PD  27-SEP-2002.
XX
PF  20-MAR-2002; 2002FR-00003459.
XX
PR  13-FEB-2001; 2001FR-00001925.
XX
PA  (MOLE-) MOLECULAR ENGINES LAB SA.
XX
PI  Telerman A, Amson R, Tuijnder M, Susini L;
XX
```

```
TCATTGTTAGACTTTGAAATTTC 26
|||||||||||||||||||||||
TCATTGTTAGACTTTGAAATTTC 42

ndard; DNA; 145 BP.

(first entry)

ession-related sequence, SEQ ID 15.

virucide; apoptotic; gene therapy; tumour suppression;
sion; apoptosis; virus resistance; viral infection; tumour;
ative disease; ds.
```

```
2002WO-FR000543.

2001FR-00001925.

CULAR ENGINES LAB.

Amson R, Tuijnder M, Susini L;

8286/05.

acid encoding a translationally controlled tumor protein,
reating, preventing and diagnosing viral, tumor or
diseases.

Page; 45pp; French.

invention relates to novel nucleic acid sequences (ABZ78294-
hich are involved in the molecular pathways of tumour
tumour reversion, apoptosis and/or virus resistance. The
e also useful for treatment or prevention of viral, tumour
enerative diseases, and also for diagnosis and prognosis of
es. Note: The sequence data for this patent is not
in the printed specification but is based on sequence
supplied by the European Patent Office

BP; 51 A; 28 C; 23 G; 43 T; 0 U; 0 Other;

larity  100.0%;  Pred. No. 0.11;  Length 145;
Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

TCATTGTTAGACTTTGAAATTTC 26
|||||||||||||||||||||||
TCATTGTTAGACTTTGAAATTTC 99

ndard; DNA; 145 BP.

(first entry)

ession-related sequence, SEQ ID 135.

virucide; apoptotic; gene therapy; tumour suppression;
sion; apoptosis; virus resistance; viral infection; tumour;
```

Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 26; Conservative 0; Mismatches 0; Indels 0; C

```
Qy    1 ATGTCATTGTTAGACTTTGAAATTTC 26
        ||||||||||||||||||||||||||
Db   74 ATGTCATTGTTAGACTTTGAAATTTC 99
```

RESULT 37
AAT19717
ID   AAT19717 standard; cDNA to mRNA; 147 BP.
XX
AC   AAT19717;
XX
DT   05-JUL-1996 (first entry)
XX
DE   Human gene signature HUMGS00791.
XX
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequenc
KW   human; cloning; mapping; non-biased library; diagnosis; detection
KW   cell typing; abnormal cell function; ss.
XX
OS   Homo sapiens.
XX
PN   WO9514772-A1.
XX
PD   01-JUN-1995.
XX
PF   11-NOV-1994;   94WO-JP001916.
XX
PR   12-NOV-1993;   93JP-00355504.
XX
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
XX
PI   Matsubara K, Okubo K;
XX
DR   WPI; 1995-206931/27.
XX
PT   Single-stranded DNA for identifying gene signatures - isolated fr
PT   directed human cDNA library that reflects relative abundance of c
PT   mRNA in specific human tissues.
XX
PS   Claim 1; Page 456; 2245pp; Japanese.
XX
CC   A single-stranded DNA (or its complementary strand or the corresp
CC   -stranded DNA) which comprises one of the 7837 "GS" sequences giv
CC   AAT19001-T26837 and which is able to hybridise to part of human g
CC   DNA, cDNA or mRNA is claimed. The GS (Gene Signature) sequences w
CC   obtained from 3'-directed cDNA libraries prepared from various hu
CC   tissues; synthesis of cDNA was initiated from the 3'-end of mRNA
CC   poly(T) as the sole primer. Since the 3'- untranslated sequence i
CC   to a particular mRNA species, almost all the 3'-oriented cDNAs hy
CC   with specific mRNAs. Each library is constructed so as to reflect
CC   accurately the relative abundance of different mRNAs in the parti
CC   tissue from which it was derived. The appearance frequency of a g
CC   in a cDNA library can be determined (esp. using primers and probe
CC   derived from the GS sequences) as a means of diagnosing abnormal
CC   function or for recognising different cell types
XX
SQ   Sequence 147 BP; 51 A; 28 C; 25 G; 43 T; 0 U; 0 Other;

Query Match   1.9%; Score 26; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 26; Conservative 0; Mismatches 0; Indels 0; G

```
Qy    1 ATGTCATTGTTAGACTTTGAAATTTC 26
        ||||||||||||||||||||||||||
Db   74 ATGTCATTGTTAGACTTTGAAATTTC 99
```

RESULT 38
AAI69032/c

204/03.

leic acid, useful for diagnosis, prognosis and treatment, ., also related vectors, transformed cell, polypeptides and

age 40; 189pp; French.

nvention relates to human oligonucleotides (ABZ08841-
e expression of the oligonucleotides is implicated in tumour
r reversion, apoptosis and/or viral resistance. The
des are useful for preventing and/or treating viral
mour development and cell degeneration (e.g. Alzheimer's
chizophrenia), especially cancer

BP; 51 A; 28 C; 23 G; 43 T; 0 U; 0 Other;

1.9%; Score 26; DB 7; Length 145;
arity 100.0%; Pred. No. 0.11;
onservative 0; Mismatches 0; Indels 0; Gaps 0;

```
CATTGTTAGACTTTGAAATTTC 26
||||||||||||||||||||||
CATTGTTAGACTTTGAAATTTC 99
```

dard; DNA; 145 BP.

(first entry)

cleotide SEQ ID 135.

suppressor; virucide; cytostatic; nootropic;
ve; neuroleptic; gene therapy; tumour suppression;
ion; apoptosis; viral resistance; viral infection;
tion; Alzheimer's disease; schizophrenia; cancer; ds.

2002FR-00003459.

2001FR-00001925.

LLAR ENGINES LAB SA.

Amson R, Tuijnder M, Susini L;

204/03.

leic acid, useful for diagnosis, prognosis and treatment, ., also related vectors, transformed cell, polypeptides and

age 58; 189pp; French.

nvention relates to human oligonucleotides (ABZ08841-
e expression of the oligonucleotides is implicated in tumour
r reversion, apoptosis and/or viral resistance. The
des are useful for preventing and/or treating viral
mour development and cell degeneration (e.g. Alzheimer's
chizophrenia), especially cancer

BP; 51 A; 28 C; 23 G; 43 T; 0 U; 0 Other;

1.9%; Score 26; DB 7; Length 145;

```
TITLE     Vogt,J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,
          NISC Comparative Sequencing Initiative
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 218485)
AUTHORS   Green,E.D.
TITLE     Direct Submission
JOURNAL   Submitted (17-JUL-2002) NIH Intramural Sequencing Center,
          Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE  3 (bases 1 to 218485)
AUTHORS   Green,E.D.
TITLE     Direct Submission
JOURNAL   Submitted (05-JUN-2003) NIH Intramural Sequencing Center,
          Grovemont Circle, Gaithersburg, MD 20877, USA
          On Jun 5, 2003 this sequence version replaced gi:26449071
COMMENT   ------- Genome Center
             Center: NIH Intramural Sequencing Center
             Center code: NISC
             Web site: http://www.nisc.nih.gov
             Contact: nisc_zoo@nhgri.nih.gov
          ------- Project Information
             Center project name: cms
             Center clone name: 145D13
          -----------------------------------------------------------
          The sequence data in this record represents an 'enhanced'
          version of a Phase 2 submission. Specifically, the indica
          order and orientation of each sequence contig has been
          established using one or more of the following: read-pair
          data from individual subclones, overlaps with neighboring
          clones, alignment with available reference sequence (e.g.
          human), and/or confirmation by PCR testing. In addition,
          the sequence assembly is based on at least 8X average
          coverage in Q20 bases and has been reviewed to rule out
          gross misassemblies, the low-quality ends of sequence
          contigs have been trimmed away, and each base is associat
          with a Phrap-derived quality score.
          --------------------- Summary Statistics
          Sequencing vector: plasmid; n/a; 100% of reads
          Chemistry: Dye-terminator Big Dye; 100% of reads
          Assembly program: Phrap; version 0.990319
          Consensus quality: 214085 bases at least Q40
          Consensus quality: 215255 bases at least Q30
          Consensus quality: 216264 bases at least Q20
          Insert size: 190000; agarose-fp
          Insert size: 216885; sum-of-contigs
          Quality coverage: 12.65x in Q20 bases; agarose-fp
          Quality coverage: 11.08x in Q20 bases; sum-of-contig
          -----------------------------------------------------------
        * NOTE: This is a 'working draft' sequence. It currently
        * consists of 17 contigs. Gaps between the contigs
        * are represented as runs of N. The order of the pieces
        * is believed to be correct as given, however the sizes
        * of the gaps between them are based on estimates that ha
        * provided by the submittor.
        * This sequence will be replaced
        * by the finished sequence as soon as it is available and
        * the accession number will be preserved.
        *
        *          1      7548: contig of 7548 bp in length
        *       7549      7648: gap of unknown length
        *       7649     31917: contig of 24269 bp in length
        *      31918     32017: gap of unknown length
        *      32018     50433: contig of 18416 bp in length
        *      50434     50533: gap of unknown length
        *      50534     95274: contig of 44741 bp in length
        *      95275     95374: gap of unknown length
        *      95375     99988: contig of 4614 bp in length
        *      99989    100088: gap of unknown length
        *     100089    109873: contig of 9785 bp in length
        *     109874    109973: gap of unknown length
        *     109974    117619: contig of 7646 bp in length
        *     117620    117719: gap of unknown length
        *     117720    128625: contig of 10906 bp in length
        *     128626    128725: gap of unknown length
        *     128726    143521: contig of 14796 bp in length
```

```
GTGCACTTGATGAGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTGGTCGGATGGT 37219

CTGGCCCTGCGCTGCGCTGGAGGAATTCTCAGCCACTGCGGCCACTGCCTCGGGCCC 720
|||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
CTGGCCCTGCGCTGCGCTGGAGGAATTCTCAGCCACTGCGGCCACTGCCTCGGGCCC 37159

CTCCGCTCTGCCAGGTGTCTGGGCGTGTTGGCCCTGCGGCCGCAGGGTCCTCCTGCCG 780
|||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
CTCCGCTCTGCCAGGTGTCTGGTCGTGTTGGCCCTGCGGCCGCAGGGTCCTCCTGCCG 37099

CGCACCCCCCTGGGCCATCAAGGCTGCCCCTTCCTCACTACTTCGGACTC 840
||||||||||||||||||||||||||||||||||||||||||||||||||
CGCACCCCCCTGGGCCATCAAGGCTGCCCCTTCCTCACTACTTCGGACTC 37039

CAGGTTCACTGAGGGCCCTGTCTCCCACAGTCGTCCCAGGCTGCGGCTCCCCT 900
|||||||||||||||||||||||||||||||||||||||||||||||||||||
CAGGTTCACTGAGGGCCCTGTCTCCCACAGTCGTCCCAGGCTGCGGCTCCCCT 36979

CAGCTCTCTGGGCACCCGGTCCCCTCTGCCCCACCCTCAGCCGCGCTCTTTGCTTCCAGA 960
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAGCTCTCTGGGCACCCGGTCCCCTCTGCCCCACCCTCAGCCGCGCTCTTTGCTTCCAGA 36919

GCCCCTCCCCTCCCTCCAGAGCCTGCCTGGGCCGTGTTCACGTGTTTTCCATCCCACATAAAT 1020
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCCCCTCCCCTCCCTCCAGAGCCTGCCTGGGCCGTGTTCACGTGTTTTCCATCCCACATAAAT 36859

GTATTCCCACTCTTATCTTACACTCCCCACGCCCACTCTCCACCTCACTAGCTC 1080
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTATTCCCACTCTTATCTTACACTCCCCACGCCCACTCTCCACCTCACTAGCTC 36799

AATCCCTGACCCTTTGAGGCCCCAGTGATTCGACTCCCCCCTGGCCACAGACCCC 1140
||||||||||||||||||||||||||||||||||||||||||||||||||||||
AATCCCTGACCCTTTGAGGCCCCAGTGATTCGACTCCCCCCTGGCCACAGACCCC 36739

GGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCCCACTTCAG 1200
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCCCACTTCAG 36679

CTAAGAGGGGCTGGACCTGGCGGCAGGAGAGCCAAAGAGACTGGGCCTAGGCCAGGAG 1260
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTAAGAGGGGCTGGACCTGGCGGCAGGAGAGCCAAAGAGACTGGGCCTAGGCCAGGAG 36619

CCAAAATGTGAGGGGCCAGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGAT 1320
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCAAAATGTGAGGGGCCAGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGAT 36559

TAAAAACAGATATTATTTTTATTATTATTGTGACAAAATGTTGATAAATGG 1373
|||||||||||||||||||||||||||||||||||||||||||||||||||
TAAAAACAGATATTATTTTTATTATTATTGTGACAAAATGTTGATAAATGG 36506
```

```
470       218485 bp    DNA     linear   HTG 05-JUN-2003
roglodytes clone RP43-145D13, WORKING DRAFT SEQUENCE, 17
ed pieces.
470.
470.4     GI:31415893
HTGS_PHASE2; HTGS_DRAFT.
roglodytes (chimpanzee)
roglodytes
yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
lia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
ages 1 to 218485)
ellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
sley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Cariaga,K.,
., Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X.,
J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
,B., Idol,J.R., Karlins,E., Kwong,P., Laric,P., Lee-Lin,S.-Q.,
pi,R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Masiello,C.,
ri,B., McDowell,J., Paguirigan,C., Pearson,R., Portnoy,M.E.,
1,A., Reddix-Dugue,N., Schandler,K., Schueler,M.G., Shah,K.,
,C., Stantripop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V.,
```

```
Db   52645  GTGCTGGCCTCGCGCTGCCTGGAGGAATTCTCAGCCACTGCGGCGAGTTCCCTCGC
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     721  CAGCTCCGCCTCTGCCAGGTGTCGGGCTGTTGGCCCTGCGGCCAGGGTCCTCCCT
Db   52705  CAGCTCCGCCTCTGCCAGGTGTCGGGCTGTTGGCCCTGCGGCCAGGGTCCTCCCT
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     781  ATCCGCACCCTCCCTCGGGCCCATCTCAAGGCTGCCCCCTTCTCACCTACTTCGG
Db   52765  ATCCGCACCCTCCCTCGGGCCCATCTCAAGGCTGCCCCCTTCTCACCTACTTCGG
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     841  TTCCAGTTCACTGAGGGCCCTGGTCTCCCACAGTCGTCCCAGGCTGCGCGGCTC
Db   52825  TTCCAGTTCACTGAGGGCCCTGGTCTCCCCAGTCGTCCCAGGCTGCGCGGCTC
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     901  CGACAGCTCTCTGGGCACCCGGTCTCCTCGCCCACCCTCAGCTGTCTTTGCTC
Db   52885  CGACAGCTCTCTGGGCACCCGGTCTCCTCGCCCACCCTCAGCTGTCTTTGCTC
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     961  CCTGCCCTCCCTCTAGAGGCTGCCTGGGCTGTTCACGTGTTTCCATCCCACAT
Db   52945  CCTGCCCTCCCTCTAGAGGCTGCCTGGGCTGTTCACGTGTTTCCATCCCACAT
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1021  ACAGTATTCCCACTCTTATCTTACAACT-CCCCACCACGCCACTCTCCACCTCACT
Db   53005  ACAGTATTCCCACTCTTATCTTACAACTCCCCCACCACGCCACTCTCCACCTCACT
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1080  CCCCATCCCTGACCCTTTGAGGCCCCAGTGATCTGGGCCAAGGATGGTCCAGAAGACCCCAC
Db   53065  CCCCATCCCTGACCCTTTGAGGCCCCAGTGATCTGGGCCAAGGATGGTCCAGAAGACCCCAC
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1140  CCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGCCAAGGATGGTCCAGAAGACCCCAC
Db   53125  CCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGCCAAGGATGGTCCAGAAGACCCCAC
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1200  GGCACTAAGAGGGGGCTGGACCTGGCGGCCAGGAAGCCAAAGAGACTGGCCTAGGCC
Db   53185  GGCACTAAGAGGGGGCTGGACCTGGCGGCCAGGAAGCCAAAGAGACTGGCCTAGGCC
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1260  GTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTG
Db   53245  GTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTG
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1320  TTTTTAAAACAGATATTATTTTTATTATTATTTGTGACAAAATGTTGATAAATGG  1
Db   53305  TTTTTAAAACAGATATTATTTTTATTATTATTTGTGACAAAGTGTTGATAAATGG  5
```

RESULT 12
AF030100
```
LOCUS       AF030100                1239 bp    mRNA    linear   ROD 30
DEFINITION  Mus musculus TWEAK mRNA, complete cds.
ACCESSION   AF030100
VERSION     AF030100.2  GI:3334855
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murin
REFERENCE   1  (bases 1 to 1239)
  AUTHORS   Chicheportiche,Y., Bourdon,P.R., Xu,H., Hsu,Y.M., Scott,H.
            Hession,C., Garcia,I. and Browning,J.L.
  TITLE     TWEAK, a new secreted ligand in the tumor necrosis factor
            that weakly induces apoptosis
  JOURNAL   J. Biol. Chem. 272 (51), 32401-32410 (1997)
  MEDLINE   98070415
  PUBMED    9405449
REFERENCE   2  (bases 1 to 1239)
  AUTHORS   Chicheportiche,Y., Bixler,S., Tizard,R. and Browning,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge
            Cambridge, MA 02142, USA
REFERENCE   3  (bases 1 to 1239)
  AUTHORS   Chicheportiche,Y., Bixler,S., Tizard,R. and Browning,J.
```

Left column (rotated):

```
         3522   143621: gap of unknown length
         3622   150832: contig of 7211 bp in length
         0833   150932: gap of unknown length
         0933   156073: contig of 5141 bp in length
         6074   156173: gap of unknown length
         6174   187450: contig of 31277 bp in length
         7451   187550: gap of unknown length
         7551   191471: contig of 3921 bp in length
         1472   191571: gap of unknown length
         1572   207263: contig of 15692 bp in length
         7264   207363: gap of unknown length
         7364   209043: contig of 1680 bp in length
         9044   209143: gap of unknown length
         9144   217567: contig of 8424 bp in length
         7568   217667: gap of unknown length
         7668   218485: contig of 818 bp in length.
```

```
Location/Qualifiers
1..218485
       /organism="Pan troglodytes"
       /mol_type="genomic DNA"
       /db_xref="taxon:9598"
       /clone="RP43-145D13"
       /clone_lib="RP43"
1..7548
       /note="assembly_fragment
       clone_end:SP6
       vector_side:left"
7649..31917
       /note="assembly_fragment"
32018..50433
       /note="assembly_fragment"
50534..95274
       /note="assembly_fragment"
95375..99988
       /note="assembly_fragment"
100089..109873
       /note="assembly_fragment"
109974..117619
       /note="assembly_fragment"
117720..128625
       /note="assembly_fragment"
128726..143521
       /note="assembly_fragment"
143622..150832
       /note="assembly_fragment"
150933..156073
       /note="assembly_fragment"
156174..187450
       /note="assembly_fragment"
187551..191471
       /note="assembly_fragment"
191572..207263
       /note="assembly_fragment"
207364..209043
       /note="assembly_fragment"
209144..217567
       /note="assembly_fragment"
217668..218485
       /note="assembly_fragment
       clone_end:T7
       vector_side:right"
```

```
arity      54.9%;   Score 754;  DB 2;  Length 218485;
           99.2%;   Pred. No. 2.9e-128;
onservative      0;  Mismatches      5;  Indels      1;  Gaps      1;
```

```
TGCACTTTGATGAGGGAAGGCTGTCTACCTGGACTTGCTGGTGGATGGT  660
     |||||||||||||||||||||||||||||||||||||||||||||
TGCACTTTGATGAGGGAAGGCTGTCTACCTGGACTTGCTGGTGGATGGT  52644
```

```
TGGGCCCTGCCTGCCTGAGGAATTCTCAGCCACTGCGCCCTCGGGCCC  720
     ||||||||||||||||||||||||||||||||||||||||||||
TGGGCCCTGCCTGCCTGAGGAATTCTCAGCCACTGCGCCCTCGGGCCC
```

```
; Submission
tted (30-JUL-2003) Cell Biology, Biogen, 12 Cambridge Center,
dge, MA 02142, USA
ice update by submitter
30, 2003 this sequence version replaced gi:2707220.
    Location/Qualifiers
    1..1239
        /organism="Mus musculus"
        /mol_type="mRNA"
        /db_xref="taxon:10090"
        /cell_type="peritoneal macrophages"
    1..750
        /note="ligand in the TNF family; secreted protein"
        /codon_start=1
        /product="TWEAK"
        /protein_id="AAC53517.2"
        /db_xref="GI:33348855"
        /translation="MAARRSQRRRGRRGEPGTALLAPLVLSLGLAIACLGLLLVVVSL
        GSWATLSAQEPSQEELTAEDRREPPLNPQTEESQDVVPFLEQLVRPRRSAPKGRKAR
        PRRAIRAHYEVHPRPGQDAGAAGVDGTVSGWEETKINSSSPLRYDRQIGEFTVIRAGL
        YLYLCQVHFDEGKAVIKLDLLIVNGVLALRCLEEFSATAASSPGPQLRLCQVSGLLPL
        RPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH"
```

```
arity     49.8%;   Score 683.4;   DB 10;   Length 1239;
77.0%;   Pred. No. 3.8e-115;
nservative   0;   Mismatches  221;   Indels  70;   Gaps   9;
```

```
Qy   :CCGCCCGTCGGAGCCAGAGGCGGAGGGGGCGCCCGGGGAGCCCGGGCCACCGCCCTG  165
          ||||||||||||||||||||||||||||  |||||||||||||||||||||||||||
Db   :CCGCCCGTCGGAGCCAGAGGCGGAGGGGGCGCCGGGGGAGCCCGGGCCACCGCCCTG   60

Qy   :TTCCCGCTCGCGCTGGGCCTGGCCCTGGCCCTGGCCCTGGCCCTCCTGCCTGCGCC  225
          |||   ||| |||||||||||||||||||||||||||||||||||||||||||||
Db   :CCCCGCTGGTGCTGAGCCTGGCCCTGGCCCTGGCCCTGGCCCTCCTGCCTGCTGTC  120

Qy   :TCAGTTTGGGGAGCCGGGATCCGTGTCCGCCCAGGAGCCTGCCCAGGAGGAGCTG  285
          ||||||| ||||||||   ||||| |||| ||||||||| |||||||||||||||
Db   :TCAGCCTGGGGAGCCGGGCAACGCTGTCTGCCCAGGAGCCTTCTCAGGAGGAGCTG  180

Qy   :CAGAGGAGGACCAGGACCCGTCGAACTGAATCCCCAGACAGAGAAAGCCAGGAT  345
          ||||||| |||  |||  ||||  ||||||||||||||||||||||||||||||
Db   :CAGAGGAGGACCGCGCCGGGAGCCCCTGAACTGAATCCCCAGACAGAGGAAAGCCAGGAT  240

Qy   :CGGCCTTCCTGAACCGACTAGTTCGGCCTTCGGCGCCTTCGGAAGTCGACTCGAACTCAAGGCCGGAAA  405
          ||||  ||| || ||||||| |||||||||||||  |||||||||||||||||||||||||||||||||
Db   :TACCTTTTCTTGAACAACTAGTCCGGCCTTCGGAGAGTGCTCCTAAAGGCCGGAAG  300

Qy   :GGGCTCGAAGAGCGATCGCAGCCCATTTGATGAGGGGAAGGCTGTCTACCTGAGC  465
          ||||||||| |||||| |||||||||||| ||||||||||| ||||||||||||||
Db   :ACCTGTACTGTCAGGTGCACTTTGATGAGGGAAAGGCTGTCATCAAGCTGGACCTC  360

Qy   :CGCAGGCAGGTGTGACGGACAGTGAGGTGGCTGGGCGGACAGTGAGGTGGCTGGGCGGAAGCCAGAATCAACAGC  525
          |||||||| |||||||||||||||||||||||||||| ||||||||||| |||||||
Db   :CACAAGCAGGTGTGATGGGACAGTGAGTGGCTGGGAAGAGACCAAAATCAACAGC  420

Qy   :GCCCTCTGCCTACAACCCCAGATCAGGGAGTTTATAGTCACCCGGGCTGGGCTC  585
          ||||| ||| |||| ||||||||||||| |||||||| |||||| |||||||||
Db   :GCCCTCTGCCTACGACCGCCAGATTGGGGAATTTACAGTCAGGGGCTGGGCTC  480

Qy   :ACCTGTACTGTCAGGTGCACTTTGATGAGGGGAAGGCTGTCTACCTGAGCTGAC  645
          ||||| |||||| |||||||||||||| ||||||||| ||||||||||||||||
Db   :ACCTGTACTGTCAGGTGCACTTTGATGAGGGAAAGGCTGTCTACCTGAGCTGAC  540

Qy   :TGGTGGATGGTGTGCTGGCCCTGGCCCTGCGGCTGCCTGAGGAATTCTCAGCCACTGCGCGCC  705
          ||||| |||| ||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db   :TGGTGAACGGTGTCCTGGCCCTGGCCCTGCGCTGCCTGGAAGAATTCTCAGCCACAGCAGACA  600

Qy   :CCCTCGGGCCCCAGTCCCGCCCTCTGCCAGGTGTCTCCAGGTGTCTGGGCGTGCCCGGCCA  765
          ||||| ||| | |||||| ||||||||||||||||||||||||||||||||||||||||||
Db   :CTCCTGGGGCCCCAGCTCCGTTTGTGCCAGGTGTCTGGGCTGTCTGGGCTGTTGCCGCTGCCGCCA  660

Qy   :CCTCCCTGCGGATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTC  825
          ||||| |||||||||||||||||||| |||||| || ||||||||||||||| ||||
Db   :CCTCCCTGCGGATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTC  825
```

```
Db   661  GGGTCTTCCTTCGGATCCGCACCCTCCCTGGGCTCATCTTAAGGCTGCCCCT
              ||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   826  ACCTACTTCGGACTCTTCCAGGTTCACTGAGGGCGCCCTCGTCTCCCACAGTCGT
              ||||||  ||||||||||||||||||||||||| |||||||||||||||||||| |
Db   721  ACCTACTTTGGACTCTTCAAGTTCACTGAGGGGCCTTGCTCTCCAGATTCCTT

Qy   886  GCTGCCCGGCTCC-----CCTCGACAGCTCTCTGGGCACCCGGTCCCCTCTGCCCC
              |||||        || ||||||||||||| |||||  |||||||||||||||||
Db   781  TTCCCTGGCTCCAGGAGCATCACCACACTCCCTACCCCACCCCACTCCTCCAC

Qy   941  CAGCGCGCTCTTTGCTCCAGACCTGCCCCCTCCCTCTAGAGCTGCTGGGCCTGTT
              | |  |||||  ||| || |||||||||||||||||| ||||| ||| || ||||
Db   841  C-GCTGTCTCCTTGGTCCAGTCCGTCTCTCC--TCAAAGGCAGCCAGAGCTTGTT

Qy   1001 GTTTTCCATCCCACATAAATACAGTATTCCCACTCTTATCTTACAAACTCCCCACC
              || || ||||                 ||  ||||| |||| |||||||||||||
Db   898  GTTTCCATTCC-------ACAGACGTATCCTTGCTCTTCTTAACATCCCATCCCA

Qy   1061 ACTCTCCACCTCACTAGCTCCCAATCCCTGACCCTTTGAGGCCCCAGTGATCTC
              ||||||| |||||||||||||||  |||||||||||||||||||||||||||
Db   951  ACTATCACCCTCACTAGCTCCCAAAGCCCCTAC------------TTATCCC

Qy   1121 CCCCCCTGGCCACAGACCCCCAGGGCATTGTGTCACTGTACTCTGTGGGCAAGG
              ||||        |||||||  ||||  | ||||| ||||||||||||||||
Db   996  CCCCACCCCACTCACCCGACCACGTGTTTATTGACTTTGTGCAC---------

Qy   1181 TCCAGAGACCCCACTTCAGGCACTAAGAGGGGGCTGGACCTGGCCGGCAGGAAGCCA
              ||||||||||||||||||||||||||  | ||| || |||||||||||||||||||
Db   1040 ------------------------------CAGGCACTGAGATGGGCTGGACCTGGTGGCAGGAAGCCA

Qy   1241 GACTGGGCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGA-AACAGACAAGC
              ||||||||   ||| | ||||||||| ||||||||| ||||| ||||||||||
Db   1083 ACCTGGGACTAGGCCAGAAGTTCCCAACTGTGAGGGGGGACAAGC

Qy   1300 CCTTGAGAATTCCCGTGGATTTTAAAACAGATATTTTTTATTATATTGTG
              ||||       |||||||||| |||||   ||||| || ||||||||| |||
Db   1143 CCTGGA-----TCCCTGTGATTTTGAAA-AGATACTATTTTTATTATTGTG

Qy   1360 ATGTTGA   1366
              |||||||
Db   1197 ATGTTAA   1203
```

```
RESULT 13
AX180714
LOCUS        AX180714                  898 bp    DNA      linear    PAT 0(
DEFINITION   Sequence 1 from Patent WO0145730.
ACCESSION    AX180714
VERSION      AX180714.1   GI:15132570
KEYWORDS     synthetic construct
SOURCE       synthetic construct
  ORGANISM   artificial sequences.
             1
REFERENCE    1
  AUTHORS    Wiley,S.R.
  TITLE      Tweak receptor
  JOURNAL    Patent: WO 0145730-A 1 28-JUN-2001;
             IMMUNEX CORPORATION (US)
FEATURES     Location/Qualifiers
  source     1..898
             /organism="synthetic construct"
             /mol_type="unassigned DNA"
             /db_xref="taxon:32630"
  CDS        52..873
             /note="unnamed protein product; human TWEAK fusio
             construct"
             /codon_start=1
             /transl_table=11
             /protein_id="CAC50004.1"
             /db_xref="GI:15132571"
             /db_xref="REMTREMBL:CAC50004"
             /translation="MATGSRTSLLLAFGLLCLPWLQEGSATSSDRMKQI
```

KIYHIENEIARIKKLIGERTRSSLGSRASLSAQEPAQEELVAEEDQDPSELNPQTEES
QDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEA
RINSSSPLRYNRQIGEFIVTRAGLIYLYCQVHFDEGKAVYLKLDLLVDGVLALRCLEE
FSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPLITYFGLFQVH"

arity   45.8%;   Score 629.2; DB 6; Length 898;
onservative 99.5%;   Pred. No. 3.3e-105;
              0; Mismatches 3; Indels 0; Gaps 0;

Qy 233 TGGGGAGCGCGGGCATCGCGTGTCCGCCAGGAGGAGCTGCCCAGGAGGAGGAGCTGGCA 291
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 251 TGGGGAGCAGCGGGCATCGCGTGTCCGCCAGGAGCTGCCCAGGAGGAGGAGCTGGCA 309

Qy 292 AGGACCAGGACCCGTCGGAACTGAATCCCAGACAGAGAAAGCCAGGATCCTGCG 351
       ||||||||||||||||||||||||||||||||||||||||||||||
Db 310 AGGACCAGGACCCGTCGGAACTGAATCCCAGACAGAAGAAGCCAGGATCCTGCG 369

Qy 352 TCCTGAACCGACTAGTTCGGCCTCGCAGAGTGCACCTAAAGGCCGGAAAACACGG 411
       |||||||||||||||||||||||||||||||||||||||||||||||||
Db 370 TCCTGAACCGACTAGTTCGGCCTCGCCGCAGAAGTGCACCTAAAGGCCGGAAAAACACGG 429

Qy 412 GAAGAGGATCGCAGCCCATTATGAAGTTCATCCACGACCTGACACGGACGGAGCG 471
       ||||||||||||||||||||||||||||||||||||||||||
Db 430 GAAGAGGCATCGCAGCCCATTATGAAGTTCATCCACGACCTGACAGAGCGAGCG 489

Qy 472 CAGGTGTGGACGGGACAGTGAGTGGCTGGGACGAAGCCAGAATCAACAGCTCCAGC 531
       ||||||||||||||||||||||||||||||||||||||||||||
Db 490 CAGGTGTGGACGGGACAGTGAGTGGCTGGGACGAAGCCAGAATCAACAGCTCCAGC 549

Qy 532 TGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGCTGGGCTCTACTAC 591
       |||||||||||||||||||||||||||||||||||||||||||||||||
Db 550 TGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGCTGGGCTCTACTAC 609

Qy 592 ACTGTCAGGTGCACTTTGATGAGGGAAGGCGTGTCTACCTGGACTTGCTG 651
       ||||||||||||||||||||||||||||||||||||||||||
Db 610 ACTGTCAGGTGCACTTTGATGAGGGAAGGCGTGTCTACCTGGACTTGCTG 669

Qy 652 ATGGTGTGCTGCGCTCCGCCCTGCGCTGCTGCCTGGAGGAATTCTCAGCCACTGCGCGGCCAGTTCC 711
       |||||||||||||||||||||||||||||||||||||||||||||||
Db 670 ATGGTGTGCTGCGCTCCGCCCTGCGCTGCTGCCTGGAGGAATTCTCAGCCACTGCGCGGCCAGTTCC 729

Qy 712 GGCCCCAGCTCCGCCTCTGCCAGGTGTCTGTTGGCCCTGCCCGCCCGCGGGTCC 771
       ||||||||||||||||||||||||||||||||||||||||
Db 730 GGCCCCAGCTCCGCCCTCTGCCAGGTGTCTGGGCCTGTTGGCCCTGCCCGCGGGTCC 789

Qy 772 TGCGGATCCGCACCCCTCCCCGTGGGCCCATCTCAAGGCTGCCCCTTCCTCACCTAC 831
       ||||||||||||||||||||||||||||||||||||||||||
Db 790 TGCGGATCCGCACCCCTCCCCGTGGGCCCATCTCAAGGCTGCCCCTTCCTCACCTAC 849

Qy 832 GACTCTTCCAGGTTCACTGAGGGGCCCTGG 865
       |||||||||||||||||||||
Db 850 GACTCTTCCAGGTTCACTGAGCGCGGCGGCCGCGG 883

source
  57   1168 bp   DNA   linear   PAT 27-AUG-2002

r necrosis factor related ligand.
57
57.1 GI:22608360
57.1 GI:22608360
1505407-A/1.
tified
tified
sified
ses 1 to 1168)
portiche,Y. and Browning,J.L.
: necrosis factor related ligand
: JP 2001505407-A 1 24-APR-2001:
INC,THE FACULTY OF MEDICINE OF THE UNIVERSITY OF GENEVA
NF family related protein
P 2001505407-A/1
1-APR-2001

FEATURES   Location/Qualifiers
source   1..1168
               /organism="unidentified"
               /mol_type="genomic DNA"
               /db_xref="taxon:32644"

ORIGIN

Query Match   44.8%;   Score 614.6; DB 6; Length 1168;
Best Local Similarity 75.8%;   Pred. No. 1.5e-102;
Matches 906; Conservative 0; Mismatches 219; Indels 70; C

Qy 178 GCGCTGGCGCTCGGCGCTGGCGCTGGCGCTTCCTGCCTCCGGCCTCCTGCCGTCGTGGTCAG
       | |||||||||||||||||||||||| |||| |||||||||||||
Db 2   GTGCTGAGCCTGGGCGCTGGCGCTGGCGCTCCTGCCTCCTGCCTGCGTGCTGGTGGTCAG

Qy 238 GGGGAGCGCGGGCATCGCGTGTCCGCCGCCCGGAAGCTGCCCAGGAGGAGCTGGTGGCAGA
       |||||||| ||| ||||| || ||||||||||||||| |||||||||||||||||
Db 62 GGGGAGCTGGGCAACGCGTCTGCCCAGGAGCCTTCTCAGGAGGAGGAGCTGACAGCAGA

Qy 298 GACCAGGACCCGTCGGAACTGAATCCCAGACAGAGAAAGCCAGGATCCTGCGCC
       ||||||||||||| ||||||||||||||||||||| |||||||||||||||
Db 122 CGCCGGGAGCCCCTGAACTGAATCCCAGACAGAGAAGCCAGGATGCGGTACC

Qy 358 CTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAACACGGC
       ||||||||||||||| ||||||||||||||||||||||||||||||||
Db 182 TTGGAACAACTAGTCCCGGCCTCGAAGAAGTGCTCCTAAAGGCCGAAGGCGGCGCC

Qy 418 AGAGCGGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACACGGACGGAGCGCA
       ||||| ||||||||||||||||||||||||||||||||||||
Db 242 CGAGCTATTGCAGCCCATTATGAGGTTCATCCTCGGCCAGACAGACAGATGGAGCACA

Qy 478 GGTGTGACGGGACAGTGAGTGGCTGGGACGAAGCCAGAATCAACAGCTCCAGCCC
       ||||||||||||||||||||||||||||| |||||||||||
Db 302 GGTGTGAATGGGACAGTGAGTGGCTGGGACGAAGAGACCAAAAATCAACAGCTCCAGCCC

Qy 538 CGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGCTGGGCTCTACTACCT
       |||||||||||||| || |||| |||| ||||||||||||||||||
Db 362 CGCTACGACCGGCCAGATTGGGGAGTTTACAGTCATCAGGGCTGGGCTCTACTACCT

Qy 598 TGTCAGGTGCACTTTGATGAGGGAAGGCGTGTCTACCTGAAGCTGGACTTGCTGGT
       |||||||||||||||||||||||||| |||||||||||||| ||||||||||
Db 422 TGTCAGGTGCACTTTGATGAGGGAAGGCGTGTCTACCTGAAGCTGGACTTGCTGGT

Qy 658 GGTGTGCTGGCGCTGCTGGCCTGGAGGAATTCTCAGCCACTGCGCGGCCAGTTCCCT
       |||||||||||||| ||| ||||| |||||||||||||||||||||
Db 482 GGTGTGCTGGCGCTGCTGGCCTGGAGGAATTCTCAGCCACCACAGCAGCAAGCTCCC

Qy 718 CCCCAGCTCCGCCTCTGCCAGGTGTCTGTTGGCCCTGTTGCCGTGCCGCGGCCAGGGTCCTC
       |||| |||||| |||||||||||||||||||||||| ||||| ||||||||||||
Db 542 CCCCAGCTCCGTTTGTGCCAGGTGTCTGGGCTGTGTTGCCGTGCCGCGGCCAGGGTCTTC

Qy 778 CGGATCCGCACCCCTCCCCCTGGGCCCCATCTCAAGGCTGCCCCTTCCTTCCTCACTACTT
       |||||||||||||| || |||||| |||||||||||||||| |||||||||
Db 602 CGGATCCGCACCCTCCCCTGGGCTCCATCTCAATGCTGCCCCTTCCTTCCTAACCTACTT

Qy 838 CTCTTCCAGGTTCACTGAGGGGCCCTGTCTCCCACAGTGTCTCCAGGGCTGCCGG
       |||||||||||||||| | |||| || | ||||||| ||||
Db 662 CTCTTCAAGTTCACTGAGGGGCCCTGCTGTCTCCCAGATTCTTAAACTTCCTGG

Qy 898 -----CCTGACAGCTCTCTGGGACACCCCGTCCCTCTGCCCCAACCCCACCCCACGCGCT
                  || | || | | || | ||||||| | |||||| |||||||
Db 722 AGGAGCATCACCACACCTCCTACCCCACTCCTCCACCCCCCTC--GCTGCT

QPALRRGRGLQAQGYGVRIQDAGVYLLYSQVLFQDVTFMGQVVSREG
IRSMPSHPDRAYNSCYSAGVFHLHQGDILSVIIPRARAKLNLSPHGTF

ORIGIN

Query Match   40.3%;   Score 553;   DB 9;   Length 1816;
Best Local Similarity   100.0%;   Pred. No. 2.8e-91;
Matches 553;   Conservative   0;   Mismatches   0;   Indels   0;

```
Qy    51 ATCCCTCGGGTCCCGGATGGGGGGCGGTGAGGCGAGGCGAGCCCGCCCCCCC.
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATCCCTCGGGTCCCGGATGGGGGGCGGTGAGGCGAGGCGAGCCCGCCCCCCC.

Qy   111 CGCCCGTCGGAGCCAGAGGCGGAGGGGGCGCCGGGGGGAGCCGGGCACCGCCTG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 CGCCCGTCGGAGCCAGAGGCGGAGGGGGCGCCGGGGGGAGCCGGGCACCGCCTG

Qy   171 CCCGCTCGCGCTGGGCCTGGCCTGGCCTGGCCTGCCTCTCCGGCCTTCCTGCTGGCC
         |||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CCCGCTCGCGCTGGGCCTGGCCTGGCCTGGCCTGCCTCTCCGGCCTTCCTGCTGGCC

Qy   231 CAGTTTGGGGAGCCGGGCATCGCGTGTCCGCCAGGAGCCTGCCCAGGAGGAGCTG
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CAGTTTGGGGAGCCGGGCATCGCGTGTCCGCCCAGGAGCCTGCCCAGGAGGAGCTG

Qy   291 AGAGGAGGACCAGGACCCGTCGAACTGAATCCCAGACAGAAGCCAGGAT
         |||||||||||||||||||||||||||||||||||||||||||||||
Db   241 AGAGGAGGACCAGGACCCGTCGAACTGAATCCCAGACAGAAAGCCAGGAT

Qy   351 GCCTTTCCTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAA
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GCCTTTCCTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAA

Qy   411 GGCTCGAAGAGCGATCGCAGCCCCATTATGAAGTTCATCCACGACCTGGACAGGAC
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GGCTCGAAGAGCGATCGCAGCCCCATTATGAAGTTCATCCACGACCTGGACAGGAC

Qy   471 GCAGGCAGTGTGACGGACAGTGAGTGGCTGGGAGAAGCCAGAATCAACAGC
         ||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCAGGCAGTGTGACGGACAGTGAGTGGCTGGGAGAAGCCAGAATCAACAGC

Qy   531 CCCTCTGCGCTACAACCGCCAGATCGGGAGTTTATAGTCACCCGGGCTGGGCTC
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CCCTCTGCGCTACAACCGCCAGATCGGGAGTTTATAGTCACCCGGGCTGGGCTC

Qy   591 CCTGTACTGTCAG  603
         |||||||||||||
Db   541 CCTGTACTGTCAG  553
```

Search completed: April 8, 2004, 02:01:50
Job time : 5276.84 secs

CCAGACCTGCCCCTCCTCTCAGAGGCTGCCTGGGCTGTTGTCACGTGTTTTCCATCCC 1012
CCAGTCCTGTCTCTCC--TCAAAGGCAGCCAGAGCTTGTTCACATGTTCCATTCC-   837

CAAATACAGTATTCCCACTCTTATCTTACAACTCCCCCACCGCGCCCACTCRCCCACCTC 1072
--ACAGAGCGTATCCTTGCTCTTCTTAACATCCCATCCCACACAACTATCCCACCTC 891

AGCTCCCCAATCCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCCCCCCCTGGCCA 1132
||||||||||||||||| |||||| |||||||
AGCTCCCAAAGCCCCTAC--------TTATCCCTGACTCCCCCACCACT 936

ACCCCCAGGGCCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCC 1192
CCGACCACGTGTTTATTGACTTTGTGCAC------ 968

TTCAGCCACTAAGAGGGGGCTGGACCTGGCGCGCAGGAAGCCAAAGAGACTGGGCCTAG 1252
--CAGGCACTGAGATGGGCTGGACCTGGCGTGCAGGAAGCCAGAGAACCTGGGACTAG 1023

AGGAGTTCCCAAATGTGAGGGGCGAGA-AACAAGACAAGTTCCTCCCTTGAGAATTC 1311
AGAAGTTCCCAACTGTGAGGGGGAAGAGCTGGGGACAAGCTCCTCCTGGA----TC 1079

GTGGATTTTAAAACAGATATTATTTTTATTATTGTGACAAAATGTTGA 1366
GTGGATTTTGAAA--AGATACTATTTTATTATTTTTGTGACAAAATGTTAA 1132

051   1816 bp   mRNA   linear   PRI 12-NOV-2002
sapiens TWE-PRIL mRNA, complete cds.
051
051.1 GI:24934973

sapiens (human)
sapiens
rota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ases 1 to 1816)
--Balade,B., Medema,J.P., Lopez-Fraga,M., Lozano,J.C.,
thoten,G.M., Picard,A., Martinez-A,C., Garcia-Sanz,J.A. and
M.
logenous hybrid mRNA encodes TWE-PRIL, a functional cell
e TWEAK-APRIL fusion protein
J. 21 (21), 5711-5720 (2002)

ases 1 to 1816)
--Balade,B., Garcia-Sanz,J.A. and Hahne,M.
-Submission
ted (21-FEB-2002) DIO, CNB, ctra de Colmenar Viejo, MADRID,
0 28049, Spain
         Location/Qualifiers
         1..1816
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /chromosome="17"
         /note="expressed in primary T lymphocytes and monocytes"
         56..1048
         /note="fusion protein; TWEAK/APRIL hybrid; TWEAK and APRIL
         are members of the TNF ligand family, corresponding to
         TNFSF12 and TNFSF13, respectively"
         /codon_start=1
         /product="TWE-PRIL"
         /protein_id="AAL90443.1"
         /db_xref="GI:24934974"
         /translation="MAARRSQRRRGRGEPGTALLVPLALGLGLALACLGLLLAVVSL
GSRASLSAQEPAQEELVAEEDQPSELNPQTEESQDPAPPLNRLVRPRRSAPKGRKTR
ARRAIAAHYEVHRPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGL
YYLYCQSSDALEAWENGERSRKRRAVLTQKQHSVLHLVPINATSKDDSDVTEVMW

GenCore version 5.1.6
opyright (c) 1993 - 2004 Compugen Ltd.

c search, using sw model

il 7, 2004, 17:30:19 ; Search time 524.669 Seconds
                         (without alignments)
                         11117.066 Million cell updates/sec

-09-245-198A-3
73
atgtcattggtagacttga.........gacaaaatgttgataaatgg 1373

ENTITY_NUC
OP 10.0 , Gapext 1.0

73863 seqs, 2124099041 residues

s satisfying chosen parameters: 6747726

gth: 0
gth: 2000000000

nimum Match 0%
ximum Match 100%
sting first 45 summaries

SUMMARIES

the number of results predicted by chance to have a
than or equal to the score of the result being printed,
d by analysis of the total score distribution.

Geneseq_29Jan04:*
geneseqn1980s:*
geneseqn1990s:*
geneseqn2000s:*
geneseqn2001as:*
geneseqn2001bs:*
geneseqn2002s:*
geneseqn2002as:*
geneseqn2003s:*
geneseqn2003bs:*
geneseqn2003cs:*
geneseqn2004s:*

ry ch	Length	DB	ID	Description
1.0	1373	2	AAV18600	Aav18600 Homo sapi
1.0	1364	6	ABK34881	Abk34881 Human cDN
1.5	1421	3	AAX56000	Aax56000 Human tum
1.2	1353	3	AAA49717	Aaa49717 Human PRO
1.2	1353	6	ABK40255	Abk40255 cDNA enco
1.6	1306	7	ACC57587	Acc57587 Polynucle
.6	1306	7	ACC57901	Acc57901 Human TWE
.6	1306	9	ADC35205	Adc35205 Human cDN
1.3	1236	7	AAV47613	Aav47613 TNF relat
1.3	1236	4	AAD04350	Aad04350 Human TRE
0.7	1030	2	AAX23424	Aax23424 Human TNR
0.8	1239	4	ADC97713	Adc97713 Murine FL
6.8	898	9	AAS03964	Aas03964 Expressio
7.8	1168	2	AAV18599	Aav18599 Mus muscu
7.8	701	2	AAX23425	Aax23425 Mouse TNR
2.4	493	8	ACH34013	Ach34013 Human end
0.5	408	7	ABX37032	Abx37032 Bovine ES
7.9	282	2	AAT22190	Aat22190 Human gen
1.1	195	6	ABK29540	Abk29540 Colon ade
7.1	412	9	ADB56326	Adb56326 Toxicity-
5.7	264	7	ABX52254	Abx52254 Bovine ES
5.1	114955	2	AAX53491	Aax53491 Human ade
4.9	3163	9	ADC87060	Adc87060 Human GPC

	24	65.4	4.8	1064	6	ABT09678	Abt09678 Hu
	25	63.8	4.6	5452	9	ADC86736	Adc86736 Hu
c	26	62.2	4.5	12733	6	ABK98631	Abk98631 Ve
c	27	62.2	4.5	12733	8	ACD13882	Acd13882 L.
c	28	62.2	4.5	12739	6	ABK98592	Abk98592 Ve
c	29	62.2	4.5	12739	8	ACD13843	Acd13843 Pl
c	30	61.4	4.5	1117	9	ADC86688	Adc86688 Hu
	31	61	4.4	1337	2	AAZ17263	Aaz17263 Hu
c	32	60.6	4.4	1000	3	AAA02484	Aaa02484 Hu
	33	60	4.4	60	6	ABN58848	Abn58848 Hu
	34	60	4.4	60	6	ABN41049	Abn41049 Hu
	35	60	4.4	60	6	ABN58593	Abn58593 Hu
	36	60	4.4	60	6	ABN58591	Abn58591 Hu
	37	60	4.4	60	6	ABN58849	Abn58849 Hu
c	38	59.6	4.3	3133	3	ADC86738	Adc86738 Hu
c	39	59.2	4.3	10732	3	AAA10594	Aaa10594 Ge
c	40	58	4.2	1218	3	AAA02488	Aaa02488 Hu
	41	56	4.1	1065	6	ABT09682	Abt09682 Hu
	42	54.8	4.0	1017	7	AAD36876	Aad36876 S.
	43	54.8	4.0	2000	7	ADA71938	Ada71938 Ri
c	44	54.8	4.0	29870	7	AAD36874	Aad36874 St
	45	54.4	4.0	600	6	ABQ52497	Abq52497 Ol

ALIGNMENTS

RESULT 1
AAV18600
ID   AAV18600 standard; cDNA; 1373 BP.
XX
AC   AAV18600;
XX
DT   21-JUL-1998 (first entry)
XX
DE   Homo sapiens tumour necrosis factor related ligand (TRELL) gene.
XX
KW   TRELL; tumour necrosis factor related ligand; tnf; treatment; can
KW   autoimmune disease; immune system; stimulation; suppression;
KW   graft rejection; ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..852
FT   /*tag= a
FT   /note= "tumour necrosis factor related ligand"
XX
PN   WO9805783-A1.
XX
PD   12-FEB-1998.
XX
PF   07-AUG-1997; 97WO-US013945.
XX
PR   07-AUG-1996; 96US-0023541P.
PR   18-OCT-1996; 96US-0028515P.
PR   18-MAR-1997; 97US-0040820P.
XX
PA   (BIOJ) BIOGEN INC.
PA   (UVGE-) UNIV GENEVA FACULTY MEDICINE.
XX
PI   Chicheportiche Y, Browning JL;
XX
DR   WPI; 1998-145619/13.
DR   P-PSDB; AAW47525.
XX
PT   Tumour necrosis factor related ligand - useful for, e.g. treating
PT   auto-immune disease and immune responses to tissue grafts.
XX
PS   Claim 2; Page 48-50; 69pp; English.
XX
CC   The sequence is that encoding human tumour necrosis factor related
CC   (TRELL). TRELL or active fragments can be included with a carrier

al compositions to treat cancer, autoimmune diseases or
uses to tissue grafts, or to stimulate or suppress the immune
s useful to screen for TRELL receptors, by labelling with a
abel and screening compositions for binding. Agents
with TRELL-receptor binding can also be screened for, can
nistered, optionally with interferon- gamma, to induce cell
at, suppress or alter immune responses (especially involving
arcinoma cells) involving a signal pathway between TRELL and
. The DNA sequence can be used in gene therapy for TRELL-
rders in mammals (especially humans), e.g. tumours,
nd inflammatory diseases or inherited genetic disorders, by
into cells, and expressing, therapeutically effective amounts
e.g. a virus comprising a gene encoding TRELL. It may also
the preparation of prepare probes for screening
netic DNAs for TRELL-encoding sequences and for antisense

3 BP; 247 A; 462 C; 394 G; 270 T; 0 U; 0 Other;

larity   100.0%;   Score 1373;   DB 2;   Length 1373;
       100.0%;   Pred. No. 1.9e-282;
Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
TCATTGTTAGACTTTGAAATTCCGCCGCCGGCTCCCCTCCCCGATCCTCGGG     60
||||||||||||||||||||||||||||||||||||||||||||||||||||
TCATTGTTAGACTTTGAAATTCCGCCGCCGGCTCCCCTCCCCGATCCTCGGG     60

CGGGATGGGGGGCGGTGAGGCAGGCACAGCCCCCCGCCCCATGCCGCCCGTCGG    120
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CGGGATGGGGGGCGGTGAGGCAGGCACAGCCCCCCGCCCCATGCCGCCCGTCGG    120

CAGAGGCGGAGGGGGGCGCCCGGGGGGAGCCGGGGCACCGCCTTGTCTCCCGTCGCG    180
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAGAGGCGGAGGGGGGCGCCCGGGGGGAGCCGGGGCACCGCCTTGTCTCCCGTCGCG    180

GGCCTGGGCCTGGCGCTGGCCTGGCGCTGGCCTGCTGGTCAGTTGGGG    240
||||||||||||||||||||||||||||||||||||||||||||||||
GGCCTGGGCCTGGCGCTGGCCTGGCGCTGGCCTGCTGGTCAGTTGGGG    240

CGGGCATCGCTGTCGCCCGCCAGGAGCCTGCCCAGGAGGAGCTGGTGGCAGAGGAGGAC    300
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGGGCATCGCTGTCGCCCGCCAGGAGCCTGCCCAGGAGGAGCTGGTGGCAGAGGAGGAC    300

GACCGGTCGAACTGAATCCCCAGACAGCAGAAGAGAAAGCCAGGATCCTGCGCCTTTCCTG    360
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GACCGGTCGAACTGAATCCCCAGACAGCAGAAGAGAAAGCCAGGATCCTGCGCCTTTCCTG    360

CGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAACACGGGCTCGAAGA    420
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAACACGGGCTCGAAGA    420

ATCGCAGCCCATTATGAAGTTCATCCAGCACCTGGACAGACAGGACGCGCAGGCAGGT    480
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ATCGCAGCCCATTATGAAGTTCATCCAGCACCTGGACAGACAGGACGCGCAGGCAGGT    480

GACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTTCCAGCCCCTCTGCGC    540
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTTCCAGCCCCTCTGCGC    540

AACCGGCAGATCGGGGAGTTTATAGTCACCCGGGGCTGGGGCTTCACTACTTGTACTGT    600
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AACCGGCAGATCGGGGAGTTTATAGTCACCCGGGGCTGGGGCTTCACTACTTGTACTGT    600

GTGCACTTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGACTGCTGGTGGTGGATGGT    660
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTGCACTTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGGCTTGCTGGTGGTGGATGT    660

CTGGCCCTGCCTCGCTGCCTGAGGAATTCCAGCCACTGCGCCACGCGCCAGTTCCCCTCGGGCCC    720
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTGGCCCTGCCTCGCTGCCTGAGGAGGAATTCTCAGCCACTGCGCCACGCGCCAGTTCCCCGGGCCC    720

CTCGGCCTCTGCCAGGTGTCTGGGCTGTTGCCCTGCGGCCAGGGTCCTCCCTGCGGG    780
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTCGGCCTCTGCCAGGTGTCTGGGCTGTTGCCCTGCGGCCAGGGTCCTCCCTGCGGG    780
```

Db   721 CAGTCCGCCGCCTGCGCCAGGTGTCTGGGCTGTTGGCCCTGCGGCCAGGGTCCTCCC
Qy   781 ATCCGCACCTCCCCTTCCTCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCCACCACTTCG
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 ATCCGCACCTCCCCTTCCTCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCCACCACTTCG
Qy   841 TTCCAGGTTCACTGAGGGGCCCTGGTCTCCCCACAGTCGTCCCAGGCTGCCGGCT
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TTCCAGGTTCACTGAGGGGCCCTGGTCTCCCCACAGTCGTCCCAGGCTGCCGGCT
Qy   901 CGACAGCTCTCTGGGCACCCGGTCCCCTGCCCCAACCCTCAGCCGCTCTTTGCT
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 CGACAGCTCTCTGGGCACCCGGTCCCCTGCCCCAACCCTCAGCCGCTCTTTGCT
Qy   961 CCTGCCCTCCCTCCCTCTAGAGGCTGCCTGGGCCTGTTCCATCCCACA
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CCTGCCCTCCCTCCCTCTAGAGGCTGCCTGGGCCTGTTCCATCCCACA
Qy   1021 ACAGTATTCCCACTCTTATCTTACAACTCCCCCACCGCCACCTCTCCACCTCACT
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACAGTATTCCCACTCTTATCTTACAACTCCCCCACCGCCACCTCTCCACCTCACT
Qy   1081 CCCAATCCCTGACCCTTTGAGGCCCCAGTGATCTTCGACTTCCCCCTGGCCACAG
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CCCAATCCCTGACCCTTTGAGGCCCCAGTGATCTTCGACTTCCCCCTGGCCACAG
Qy   1141 CAGGGCATTGTGTTCACTGTACTGTGTGGGCAAGGATGGGTCCAGAAGACCCCAC
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CAGGGCATTGTGTTCACTGTACTGTGTGGGCAAGGATGGGTCCAGAAGACCCCAC
Qy   1201 GCACTAAGAGGGGCTGGACCTGGCCGGCAGGAAGCCAAAGAGACTGGGCTTAGGCC
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GCACTAAGAGGGGCTGGACCTGGCCGGCAGGAAGCCAAAGAGACTGGGCCTAGGCC
Qy   1261 TTCCCAAATGTGAGGGCGGAGAAACAAGACAAGCTCCTCCTTGAGAATTCCCTG
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 TTCCCAAATGTGAGGGCGGAGAAACAAGACAAGCTCCTCCTTGAGAATTCCCTG
Qy   1321 TTTTAAACAGATATTATTTTTATTATTGTGACAAAATGTTGATAAATGG 1
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 TTTTAAACAGATATTATTTTTATTATTGTGACAAAATGTTGATAAATGG 1

RESULT 2
ABK34881
ID   ABK34881 standard; cDNA; 1364 BP.
XX
AC   ABK34881;
XX
DT   08-MAY-2002 (first entry)
XX
DE   Human cDNA encoding secreted protein #19.
XX
KW   Human; secreted protein; gene; ss; nutritional supplement; haemo
KW   viral infection; bacterial infection; fungal infection; diabetes
KW   autoimmune disorder; rheumatoid arthritis; multiple sclerosis; t
KW   autoimmune thyroiditis; allergic reaction; neurodegenerative dis
KW   Alzheimer's disease; Parkinson's disease; liver fibrosis; cancer
KW   coagulation disorder; inflammatory disorder; Crohn's disease; in
KW   tissue regeneration; wound healing; burn; haematopoiesis;
KW   myeloid cell deficiency; lymphoid cell deficiency.
XX
OS   Homo sapiens.
XX
PN   WO200177288-A2.
XX
PD   18-OCT-2001.
XX
PF   29-MAR-2001; 2001WO-US010224.
XX
PR   06-APR-2000; 2000US-0195582P.
XX
PA   (GEMY) GENETICS INST INC.

k HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;
raham JR;

21/23.

nd ninety two polynucleotides derived from a variety of
ources which encode secreted proteins, useful for treating
ncies and disorders such as autoimmune disorders.

82; 372pp; English.

relates to 592 polynucleotides which have been derived from
uman tissue sources and which encode novel secreted
polynucleotides can be used as probes for the
and isolation of full length cDNA and genomic DNA. The
s and proteins can also be used as nutritional supplements.
re useful in the treatment of various immune deficiencies
such as viral infections, bacterial infections, fungal
toimmune disorders (e.g. rheumatoid arthritis, multiple
oimmune thyroiditis and diabetes) and allergic reactions
(e.g. asthma). They are also useful for treating
ive diseases (e.g. Alzheimer's disease, Parkinson's
r fibrosis, coagulation disorders (e.g. haemophilia),
disorders (e.g. Crohn's disease) and tumours. They are also
sue regeneration, for wound healing and in the treatment of
ns and ulcers. The proteins are also useful for regulating
and for treating myeloid or lymphoid cell deficiencies.
4863-ABK35454 represent polynucleotides of the invention

BP; 246 A; 461 C; 394 G; 263 T; 0 U; 0 Other;

rity   98.0%;   Score 1345.8; DB 6;   Length 1364;
nservative 99.9%; Pred. No. 1.2e-276;
            0; Mismatches   2;   Indels   0; Gaps   0;

| | | | | | | | | | | | | | | | | |
Qy ...CCGCCGGGTCCCCCTCCCCCGATCCTCGGGTCCGGATGGGGGGGCCGTGAGG 84
    | | | | | | | | | | | | | | | | | |
Db ...CCGCCGGGTCCCCCTCCCCCGATCCTCGGGTCCGGATGGGGGGGCCGTGAGG 64

Qy ...ACAGCCCCCCCGCCCCCATGGCCGCCCGTCGGAGCCAGAGGCGGAGGCGCCCGG 144
    | | | | | | | | | | | | | | | | | |
Db ...ACAGCCCCCCCGCCCCCATGGCCGCCCGTCGGAGCCAGAGGCGGAGGCGCCCGG 124

Qy ...GCCGGGCACCGCCCCTGCTGTCCCGCTCGCGCTGGGCCTGGCCTGGCC 204
    | | | | | | | | | | | | | | | | | |
Db ...GCCGGGCACCGCCCCTGCTGTCCCGCTCGCGCTGGGCCTGGCCTGGCC 184

Qy ...CGGCCTCTGCTGCCGTGTGTCAGTTGGGAGCCGGAGCATCGCTGTCCGCCCAG 264
    | | | | | | | | | | | | | | | | | |
Db ...CGGCCTCTGCTGCCGTGTGTCAGTTGGGAGCCGGAGCATCGCTGTCCGCCCAG 244

Qy ...TGCCCAGGAGGAGAGCTGGTGGCAGAGGAGGAGGACCCAGGACCCGTCCGCCCAGA 324
    | | | | | | | | | | | | | | | | | |
Db ...TGCCCAGGAGGAGAGCTGGTGGCAGAGGAGGAGGACCCAGGACCCGTCCGCCCAGA 304

Qy ...AGAGAAAGCCAGGATCCTGCGCCTTTCCTGAACCGACTAGTTCGGCCTCGCAGA 384
    | | | | | | | | | | | | | | | | | |
Db ...AGAGAAAGCCAGGATCCTGCGCCTTTCCTGAACCGACTAGTTCGGCCTCGCAGA 364

Qy ...ACCTAAAGGCCGGAAAAACACGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTT 444
    | | | | | | | | | | | | | | | | | |
Db ...ACCTAAAGGCCGGAAAAACACGGGCTCGAAGAGCCCATTATGAAGTT 424

Qy ...ACGACCTGACAGGAGGACGAGGCAGGCAGCAGGTGTGACGGGACAGTGAGTGGCTGG 504
    | | | | | | | | | | | | | | | | | |
Db ...ACGACCTGACAGGAGGACGAGGCAGGCAGGTGTGACGGGACAGTGAGTGGCTGG 484

Qy ...AGCCAGAATCAACAGCTCCAGCCCTCTGCGCTACAACCGCCAGATCGGGGAGTTT 564
    | | | | | | | | | | | | | | | | | |
Db ...AGCCAGAATCAACAGCTCCAGCCCTCTGCGCTACAACCGCCAGATCGGGGAGTTT 544

Qy ...CACCCGGGCTGGGCTTCTACTACCTGTACTGTCAGGTGCACTTTGATGAGGGGAAG 624

Db 545 ATAGTCACCCGGGCTGGGCTTCTACTACCTGTACTGTCAGGTGCACTTTGATGAGGG
         | | | | | | | | | | | | | | | | | |
Qy   625 GCTGTCTACCTGAAGCTGGACTTGCTGGTGGATGGTGTGCTGGCCCTGCGCGCTGCCT
         | | | | | | | | | | | | | | | | | |
Db   605 GCTGTCTACCTGAAGCTGGACTTGCTGGTGGATGGTGTGCTGGCCCTGCGCGCTGCCT

Qy   685 GAATTCTCAGCCATTGCGGCCAGTTCCCTGGGCCCCAGCTCCGCCTCTGCCAGGT
         | | | | | | | | | | | | | | | | | |
Db   665 GAATTCTCAGCCACTGCGGCGGAGTTCCCTGGGCCCCAGCTCCGCCTCTGCCAGGT

Qy   745 GGGCTGTTGGCCGTGCGGCCAGGGTCCTCCCTCCCTGGGATCCGCACCCTCCCTGGGC
         | | | | | | | | | | | | | | | | | |
Db   725 GGGCTGTTGGCCCTGCGGCCAGGGTCCTCCCTCCCTGCGGATCCGCACCCTCCCTGGGC

Qy   805 CTCAAGGCTGCCCCTTCTCACCTACTTCGGACTCTTCCAGGTTCACTGAGGGGCC
         | | | | | | | | | | | | | | | | | |
Db   785 CTCAAGGCTGCCCCCTTCTCCACCTACTTCGGACTCTTCCAGGTTCACTGAGGGGCC

Qy   865 GTCTCCCCACAGTCGTCCCAGGCGTGCGGCTCCCCTGACAGCTCTCTGGGCACCC
         | | | | | | | | | | | | | | | | | |
Db   845 GTCTCCCCGCAGTCGTCCCAGGCGTGCGGCTCCCCTGACAGCTCTCTGGGCACCC

Qy   925 CCCCTCGCCCCACCCTCAGCCGCGCTCTTTGCTCCAGACTGCCCCTCCCTCTCAGAGG
         | | | | | | | | | | | | | | | | | |
Db   905 CCCTCGCCCCACCCTCAGCCGCGCTCTTTGCTCCAGACCTGCCCCTCCCTCTCAGAGG

Qy   985 CTGGGCCTGTTCAGTGTTTCATCCCACATAAATACAGTATTCCCACTCTTATC
         | | | | | | | | | | | | | | | | | |
Db   965 CTGGGCCTGTTCAGTGTTTTCATCCCACATAAATACAGTATTCCCACTCTTATC

Qy 1045 AACTCCCCACCGCCACTCTCCACTCTCCACTAGCTCCCCAATCCCTGACCCTTTGA
         | | | | | | | | | | | | | | | | | |
Db 1025 AACTCCCCACCGCCACTCTCCACTCTCCACTAGCTCCCCAATCCCTGACCCTTTGA

Qy 1105 CCCAGTGATCTCGACTCCCCCTGCCACAGACCCCAGGGCATTGTGTTCACTGT
         | | | | | | | | | | | | | | | | | |
Db 1085 CCCAGTGATCTCGACTCCCCCTGCCACAGACCCCAGGGCATTGTGTTCACTGT

Qy 1165 TGTGGGCAAGGATGGGTCAGAGAGACCCCACTTCAGGCACTAAGAGGGGCTGGACC
         | | | | | | | | | | | | | | | | | |
Db 1145 TGTGGGCAAGGATGGGTCAGAGAGACCCCACTTCAGGCACTAAGAGGGGCTGGACC

Qy 1225 GGCAGGAAGCCAAAGAGACTGGGCTTAGGCCAGGAGTTCCCAAATGTGAGGGGCGA
         | | | | | | | | | | | | | | | | | |
Db 1205 GGCAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGA

Qy 1285 CAAGACAAGCTCCTCCTTGAGAATTCCCTGTGAGTTTTTAAAACAGATATTATTT
         | | | | | | | | | | | | | | | | | |
Db 1265 CAAGACAAGCTCCTCCCTGAGAATTCCCTGTGGATTTTTAAAACAGATATTATTT

Qy 1345 TATTATTGTGACAAAATGTTGATAAATGG 1373
         | | | | | | | | | | | | | | | | | |
Db 1325 TATTATTGTGACAAAATGTTGATAAATGG 1353

RESULT 3
AAX56000
ID   AAX56000 standard; DNA; 1421 BP.
XX
AC   AAX56000;
XX
DT   15-JUL-1999 (first entry)
XX
DE   Human tumour necrosis factor Apo-3 ligand polynucleotide sequence.
XX
KW   Human; tumour necrosis factor; Apo-3 ligand; lymphotoxin; apoptos:
KW   NF-kappaB-dependent transcription; JNK/SAPK-dependent response; c&
KW   ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers

92..841
/*tag= a
/product= "Apo-3 ligand"

98WO-US021407.
97US-0062037P.
97US-0069862P.

NTECH INC.

Marsters SA, Pitti R;

7982/24.
9369.

o3- ligand (a tumor necrosis factor) homologue.

g 1; 74pp; English.

sequence encodes a human tumour necrosis factor (TNF) and
homologue designated Apo-3 ligand. Apo-3 ligand has
ctivity. Apo-3 ligand can be used to induce apoptosis in
cer cells, to induce NF-kappaB-dependent transcription and
K/SAPK-dependent responses in mammalian cells

L BP; 281 A; 464 C; 404 G; 272 T; 0 U; 0 Other;

larity 96.5%;   Score 1325;   DB 2;   Length 1421;
        98.5%;   Pred. No. 3.1e-272;
onservative   0;   Mismatches 20;   Indels   0;   Gaps   0;

Qy AAATTCCGCCCGCCGGCTCCCCCTCCCCGATCCCCCGGTCCCGGGATGGGGGGG   76
       |||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db AGATCCCCTCGACCTCGACCCCAGCGTCCGCGATCCCTCGGTCCCGGGATGGGGGGG   62

Qy TGAGGCAGGCAGCAGCCCCCCGCCCCCCATGGCGCCCCGTCGGAGCCAGAGAGGCGGAGGG   136
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db TGAGGCAGGCACAGCCCCCCGCCCCCCATGGCGCCCCGTCGGAGCCAGAGAGGCGGAGGG   122

Qy CCGGGGGGAGGACGCCGGGCACCGCCCTCGTTCCCGTCGCCGCTGGGCCTGGGCCTGG   196
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db CCGGGGGGAGGACGCCGGGCACCGCCCTCGTTCCCGTCGCCGCTGGGCCTGGGCCTGG   182

Qy GGGCCTGCCTCGGCCTCTCCTGCTGGCCGTGGTCAGTTTGGGGAGCCGCGGGCATGCGTGT   256
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db GGGCCTGCCTCGGCCTCTCCTGCTGGCCGTGGTCAGTTTGGGGAGCCGCGGGCATGCGTGT   242

Qy CCAGGAGAGCCTGCCCAGGAGGAGCTGGTGCAGAGAGGACCAGGAGACCCGGAAC   316
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db CCAGGAGAGCCTGCCCAGGAGGAGCTGGTGCAGAGAGGACCAGGAGACCCGGAAC   302

Qy ATCCCCAGACAGAAGAAGCCAGGATCCTGCGCCTTCCTGAACCGACTAGTTCGGC   376
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ATCCCCAGACAGAAGAAGCCAGGATCCTGCGCCTTCCTGAACCGACTAGTTCGGC   362

Qy GCAGAAGTGCACCTAAAGGCCGGAAACACGGGCTCGAAGAGCCGATCGCAGCCCATT   436
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db GCAGAAGTGCACCTAAAGGCCGGAAACACGGGCTCGAAGAGCCGATCGCAGCCCATT   422

Qy AAGTTCATCCACGACCTGGACAGGACGACGGCCAGGCAGGTGTGACGGGACAGTGA   496
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db AAGTTCATCCACGACCTGGACAGGACGACGGCCAGGCAGGTGTGACGGGACAGTGA   482

Qy GCTGGGAGGAAGCCAGAATCAACAGTCCTCTGCGCTACAACCGCCAGATCG   556
       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db GCTGGGAGGAAGCCAGAATCAACAGTCCTCTGCGCTACAACCGCCAGATCG   542

Qy AGTTTATAGTCACCCGGGCTGGGCTCTACTACCTGTACTGTCAGGTGCACTTTGATG   616
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db GGGAGTTTATAGTCACCCGGGCTGGGCTCTACTACCTGTACTGTCAGGTGCACTT   543

Qy AGGGGAAGCTGTCTACCTGAAGCTGGACTTGCTGGTGGATGGTGTGCTGGCCCCT   617
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db AGGGGAAGCTGTCTACCTGAAGCTGGACTTGCTGGTGGATGGTGTGCTGGCCCCT   603

Qy GCCTGGAGGAATTCTCAGCCACTGCGGCCCAGTCCCTCGGGCCCCAGTTCCGCCT   677
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db GCCTGGAGGAATTCTCAGCCACTGCGGCCCAGTCCCTCGGGCCCCAGTTCCGCCT   663

Qy AGGTGTCTGGCGTTGGCCTGGCCCGGCCAGGGTCCTCCCTCGCGGATCCGCGACCCCT   737
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db AGGTGTCTGGCGTTGGCCTGGCCCGGCCAGGGTCCTCCCTCGCGGATCCGCGACCCCT   723

Qy GGGCCCATCTCAAGGCTGCCCCCTTCCTCACCTACTTCGGGACTCTTCCAGGTTCA   797
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db GGGCCCATCTCAAGGCTGCCCCCTTCCTCACCTACTTCGGGACTCTTCCAGGTTCA   783

Qy GGGCCCTGGTCTCCCCACAGTCGTCCCAGGCTGCGGCGTCCCCTGCAGCAGCTCTC   857
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db GGGCCCTGGTCTCCCCAGTCGTCCCCGGCTGCCGGCTGCCGGCTCCCCTGCAGCTCTC   843

Qy ACCCGGTCCCCTGCCCAGCCGTCTTCGTGCCAGACCTGCCCCTCCI   917
       |||||||||||||||||||||||||||||||||||||||||||||||
Db ACCCGGTCCCCTGCCCAGCCGTCTTTGCTGCCAGACCTGCCCCTCCI   903

Qy GAGGCTGCCTGGCCTGTTCAGTGTTTTCCATCCCAATAAATACAGTATTCCC   977
       |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db GAGGCTGCCTGGCCTGTTCAGTGTTTTCCATCCCAATAAATACAGTATTCCC   963

Qy TATCTTACAACTCCCCACCGCCCACTCTCCACCTCACTAGCTCCCAATCCCTG.   1037
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db TATCTTACAACTCCCCACCGCCCACTCTCCACCTCACTAGCTCCCCAATCCCTG: 1023

Qy TTGAGGCCCCCAGTGATCTCGACTCCCCCCCTGGCCACAGACCCCAGGGCATTGT   1097
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db TTGAGGCCCCCAGTGATCTGGACTCCCCCCCTGGCCACAGACCCCCAGGTCATTGT   1083

Qy CTGTACTCTGTGGGCAAGGATGGTCAGAAGACCCCACTTCAGGCACTAAGAGAG   1157
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db CTGTACTCTGTGGGCAAGGATGGTCCAGAAGACCCCACTTCAGGCACTAAGAGAG   1143

Qy GACCTGGCGGCAGGAAGAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGT   1217
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db GACCTGGCGGCAGGAAGAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGT   1203

Qy GCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAG   1277
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db GCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAG   1263

Qy ATTTTTATTATTATTGTGACAAAATGTTGATAAATGG   1337
       |||||||||||||||||||||||||||||||||||||
Db ATTTTTATTATTATTGTGACAAAATGTTGATAAATGG   1323

Location/Qualifiers
58. .807
/*tag= a
58. .177
/*tag= b
178. .804
/*tag= c

2.

99WO-US028565.

98US-0113296P.
99WO-US005028.
99US-0130232P.
99US-0131445P.
99US-0134287P.
99US-0144758P.
99US-0145698P.
99WO-US021090.
99WO-US021547.

ECH INC.

Goddard A, Godowski PJ, Gurney AL, Marsters SA;
tti RM, Wood WI;

568/38.
838.

ion to inhibit neoplastic cell growth or for treating tumor
rises polypeptides PRO179, PRO207, PRO320, PRO219, PRO221,
3, PRO301, PRO221, PRO526, PRO362, PRO356, PRO509 or PRO866.

3; 172pp; English.

equence is that of cDNA clone DNA30879-1152 (ATCC 209358)
PRO207 (see AAY95338), which shows homology to several
tumour necrosis factor family, especially human cDNA
3.4%). The cDNA was identified in a foetal kidney cDNA
ing identification of an expressed sequence tag with
uman Apo-2 ligand. A claimed method for inhibiting the
mour cell comprises exposing the tumor cell to PRO179,
, PRO219, PRO221, PRO224, PRO328, PRO301, PRO526, PRO362,
or PRO866 (see AAY95337-49), their agonists or chimeric
ncorporating them. The tumour is especially a cancer
breast, ovarian, renal, colorectal, uterine, prostate,
and central nervous system cancer, melanoma and leukaemia.
encoding PRO179 etc. are used in the recombinant production
our polypeptides

BP: 257 A; 443 C; 389 G; 264 T; 0 U; 0 Other;

rity   96.2%;   Score 1320.2;   DB 3;   Length 1353;
        99.8%;   Pred. No. 3.3e-271;
nservative   0;   Mismatches   3;   Indels   0;   Gaps   0;

181	Db	GTCAGTTTGGGAGCCGGCATCGCTCGTCCCCCAGGAGGAGCTGCCCAGGAGGAGCT
289	Qy	GCAGAGGAGGAGGACCAGGACCCGTCGGAACTGAATCCCCAGACAGAGAAAGCCAGGA
241	Db	GCAGAGGAGGAGGACCAGGACCCGTCGGAACTGAATCCCCAGACAGAGAAAGCCAGGA
349	Qy	GCGCCTTCCTGAACCGACTAGTTCGGCCTCGCCAGAAGTGCACCTAAAGGCCGGAA
301	Db	GCGCCTTCCTGAACCGACTAGTTCGGCCTCGCCAGAAGTGCACCTAAAGGCCGGAA
409	Qy	CGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGA
361	Db	CGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGA
469	Qy	GCGCAGGCAGGTGTGGACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAG
421	Db	GCGCAGGCAGGTGTGGACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAG
529	Qy	AGCCCTCGCCTACAACCGCCAGATCGGGAGTTTATAGTCACCCGGGCTCGGGCT
481	Db	AGCCCTCGCCTACAACCGCCAGATCGGGAGTTTATAGTCACCCGGGCTCGGGCT
589	Qy	TACCTGTACTGTCAGGTGCACTTTGATGAGGGGAAGGCTGTCTACCTGAGCTGGA
541	Db	TACCTGTACTGTCAGGTGCACTTTGATGAGGGGAAGGCTGTCTACCTGAGCTGGA
649	Qy	CTGGTGGATGGTGTGCTGGCCCTGCCTGGCTCGCCTCGGAGGAATTCTCAGCCACTGCGGC
601	Db	CTGGTGGATGGTGTGCTGGCCCTGCCTGGCTCGCCTCGGAGGAATTCTCAGCCACTGCGGC
709	Qy	TCCCTCGGCCCCAGCTCCGCCTCTGCCAGGTGTCTGGGCTGTTGGCCCTGCGGCC
661	Db	TCCCCCGGGCCCCAGCTCCGCCTCTGCCAGGTGTCTGGGCTGTTGGCCCTGCGGCC
769	Qy	TCCTCCTGGGATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCT
721	Db	TCCTCCTGCGGATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCT
829	Qy	TACTTCGGACTCTTCCAGTTCACTGAGGGGCCCTGGTCTCCCCACAGTCGTCCCA
781	Db	TACTTCGGACTCTTCCAGTTCACTGAGGGGCCCTGGTCTCCCCAGTCGTCCCA
889	Qy	GCCGGCTCCCCTGCGACAGCTCTCGGCACCCGGTTCCCCTGCCCCTGCCCACCCTCAGC
841	Db	GCCGGCTCCCCTGCGACAGCTCTCGGCACCCGGTTCCCCTGCCCCTGCCCACCCTCAGC
949	Qy	CTTTGCTCAGACCTGCCCCTCCCTCTAGAGGCTGCCTGGGCCTGTTCAGTGTTT
901	Db	CTTTGCTCAGACCTGCCCCTCCCTCTAGAGGCTGCCTGGGCCTGTTCAGTGTTT
1009	Qy	TCCCACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCCACTC
961	Db	TCCCACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCCACTC
1069	Qy	CCTCACTAGCTCCCAATCCTGACCCTTTGAGGCCCCCAGTGATCTGACTCCCC
1021	Db	CCTCACTAGCTCCCAATCCTGACCCTTTGAGGCCCCCAGTGATCTGACTCCCC
1129	Qy	GCCACAGACCCCCAGGCATTGTGTTCACTGTACTCGTGTGGGCAAGGATGGGTCCA
1081	Db	GCCACAGACCCCCAGGTCATTGTGTTCACTGTACTCGTGTGGGCAAGGATGGGTCCA
1189	Qy	ACCCCACTTCAGGCACTAAGAGGGGCTGACCTGCGCGCAGGAAGCCAAAGAGACT
1141	Db	ACCCCACTTCAGGCACTAAGAGGGGCTGACCTGCGCGCAGGAAGCCAAAGAGACT
1249	Qy	CTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCTCCCCTTG
1201	Db	CTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCTCCCCTTG
1309	Qy	TTCCCTGTGGATTTTAAAACAGATATTATTTTATTATTTGTGACAAAATGTT
1261	Db	TTCCCTGTGGATTTTAAAACAGATATTATTTTATTATTTGTGACAAAATGTT

CCTCGGGTCCCGGGATGGGGGGGCCGGTGAGGCAGGCACGCCCCCGCCCCCATG 108
CCTCGGGTCCCGGGATGGGGGGGCCGGTGAGGCAGGCAGCACCCCCGCCCCCATG 60

CCGTCGGAGGCCAGAGGCCGGAGGGGGCGCCGGGGGGGGAGCCGGGCACCGGCCGCCTGCTG 168
CCGTCGGAGGCCAGAGGCCGGAGGGGGCGCCGGGGGGGGAGCCGGGCACCGGCCGCCTGCTG 120

GCTCGCCGCTGGGCCTGGGCCTGGGCCTGCCCTGCGCCTGCGCCTCGGCCTTCGCCTGGCCGTG 228
GCTCGCCGCTGGGCCTGGGCCTGGGCCTGCCCTGCGCCTGCGCCTCGGCCTTCGCCTGGCCGTG 180

TTTGGGAGCCGGCATCGCTCGTCCGCCCAGGAGCTGCCCAGGAGGAGCTGGGTG 288

GG 1373
||
GG 1325

ndard; cDNA; 1353 BP.

(first entry)

g human PRO207 polypeptide.

benign tumour; malignant tumour; lymphoid malignancy;
euronal disorder; stromal disorder; blastocoelic disorder;
disorder; immune disorder; angiogenic disorder;
; cytostatic; neuroprotective; gene; ss.

A1.

2000WO-US003565.

99WO-US005028.
99US-0123972P.
99US-0133459P.
99WO-US012252.
99US-0140650P.
99US-0140653P.
99US-0144758P.
99US-0145698P.
99US-0146222P.
99US-0149395P.
99US-0151689P.
99WO-US020111.
99WO-US021090.
99WO-US028313.
99WO-US028301.
2000WO-US000219.

TECH INC.

Goddard A, Godowski PJ,   Gurney AL,   Hillan KJ;
Pan J,   Pitti RM,   Roy MA,   Smith V,   Stone DM;
Wood WI;

567/26.
129.

ucleic acids encoding PRO polypeptides, useful for treating
ignant tumors, leukemias and lymphoid malignancies,
angiogenic and immunologic disorders.

3; 302pp; English.

nvention relates to the isolation of novel human PRO
and the polynucleotide sequences encoding them. The PRO
agonists, antagonists or anti-PRO antibodies are useful for
gn or malignant tumours (e.g. renal, kidney, bladder,
leukaemias and lymphoid malignancies, other disorders such
glial, astrocytal, hypothalamic, glandular, macrophagal,
blastocoelic disorders, inflammatory, immune and angiogenic
e polynucleotide sequences are also useful in gene therapy.
0288 encode for the human PRO polypeptides of the invention

BP; 257 A; 443 C; 389 G; 264 T; 0 U; 0 Other;

Query Match   96.2%;   Score 1320.2; DB 6;   Length 1353;
Best Local Similarity 99.8%;   Pred. No. 3.3e-271;
Matches 1322; Conservative   0; Mismatches   3;   Indels   0;

Qy   49 CGATCCCTCGGTCCCGGGATGGGGGGGCGGTGAGGCAGGCAGACAGCCCCCGCCC
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CGATCCCTCGGTCCCGGGATGGGGGGGCGGTGAGGCAGGCAGACAGCCCCCGCCC

Qy   109 GCCGCCCGTCGGAGCCAGAGGCGGAGGGGGCGCCGGGGGGAGCCGGGCACCGCCC
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GCCGCCCGTCGGAGCCAGAGGCGGAGGGGGCGCCGGGGGGAGCCGGGCACCGCCC

Qy   169 GTCCCGCTCGCTCGGGCCTGGCCTGGCCTGCCTCTCGGCCTCCTTCGGCCTCTGG
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTCCCGCTCGCTCGGGCCTGGCCTGGCCTGCCTCTCGGCCTCCTTCGGCCTCTGG

Qy   229 GTCAGTTTGGGGAGCCGGGCATCGCTGCTCGCCCAGGAGCCTGCCCAGGAGGAGC
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GTCAGTTTGGGGAGCCGGGCATCGCTGCTCGCCCAGGAGCCTGCCCAGGAGGAGC

Qy   289 GCAGAGGAGGAACCAGGACCCGTCGGAACTGAATCCCAGACAGAGAAAGCCAGG
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GCAGAGGAGGAACCAGGACCCGTCGGAACTGAATCCCAGACAGAGAAAGCCAGG

Qy   349 GCGCCTTCTGAACCGACTAGTTCGGCCTTCGCGAGAAGTGCACCTAAAGGCCGGA
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GCGCCTTCTGAACCGACTAGTTCGGCCTTCGCGAGAAGTGCACCTAAAGGCCGGA

Qy   409 CGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGACAGG.
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGACAGG.

Qy   469 GCGCAGGCAGGTGTGACGGGACAGTGAGTGGCTGGGACGGAAGCCAGAATCAACA
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCGCAGGCAGGTGTGACGGGACAGTGAGTGGCTGGGACGGAAGCCAGAATCAACA

Qy   529 AGCCCCTTCGCCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGC
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AGCCCCTTCGCCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGC

Qy   589 TACCTGTACTGTCAGGTGCACTTTGATGAGGGAAGGCTGTGTACCTGAAGCTGG.
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TACCTGTACTGTCAGGTGCACTTTGATGAGGGAAGGCTGTGTACCTGAAGCTGG.

Qy   649 CTGGTGGATGGTGTGCTGCTGGCCCTGCCCTGCCCTGGAGGAATTCTCAGCCACTGCGGC
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CTGGTGGATGGTGTGCTGCTGGCCCTGCCCTGCCCTGGAGGAATTCTCAGCCACTGCGGC

Qy   709 TCCCTCGGGCCCCAGCTGCCCTTCGCCAGGTGTCTGGGCGTGTTGGCCCTGGCGC.
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TCCCTCGGGCCCCAGCTCCGCCTTCGCCAGGTGTCTGGGCGTGTTGGCCCTGGCGC.

Qy   769 TCCTCCCTGCGGATCCGCACCCTCCCCCTCCCGGCCCATCTCAAGGCTGCCCCCTTCC.
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 TCCTCCCTGCGGATCCGCACCCTCCCCCTCCCGGCCCATCTCAAGGCTGCCCCCTTCC.

Qy   829 TACTTCGGACTCTTCCAGGTTCACTGAGGGCCCTGGTCTCCCCACAGTCGTCCCA.
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 TACTTCGGACTCTTCCAGGTTCACTGAGGGCCCTGGTCTCCCCGCAGTCGTCCCA.

Qy   889 GCCGGCTCCCCTGCGACAGCTCTCTGCGACACCCGGCACCCGGTCCCCTCTGCCCACCCTCAGC.
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GCCGGCTCCCCTGCGACAGCTCTCTGCGACACCCGGCACCCGGTCCCCTCTGCCCACCCTCAGC.

Qy   949 CTTTGCTCCAGACCTGCCCCTCCTTCTAGAGGCTGCCTGGGCCGTGTTCACGTGTTT.
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 CTTTGCTCCAGACCTGCCCCTCCTTCTAGAGGCTGCCTGGGCCGTGTTCACGTGTTT.

Qy 1009 TCCCATAAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCCACTC.
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TCCCACATAAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCCACTC.

Query Match   93.6%;   Score 1285; DB 7;   Length 1306;
Best Local Similarity 100.0%; Pred. No. 1e-263;
Matches 1285; Conservative 0; Mismatches 0; Indels 0; G

```
Qy     89 CACAGCCCCGCGCCCCATGGCGCCCGTCGGAGCCAGAGGCGGAGGGGGCGCCGG 1128
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CACAGCCCCGCGCCCCATGGCGCCCGTCGGAGCCAGAGGCGGAGGGGGCGCCGG 1080

Qy    149 AGCCGGCACCGACCCTGTGTCCGCTCGCGTCGGCCTGGCGTCGGCGCTGGCC 1188
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGCCGGCACCGACCCTGTGTCCGCTCGCGTCGGCCTGGCGTCGGCGCTGGCC 1140

Qy    209 TCGGCCTCCTGCTGCCGTGGTCAGTTGGGGAGGAGCCGGGCATCGTGTCCGCCCAG 1248
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TCGGCCTCCTGCTGCCGTGGTCAGTTGGGGAGGAGCCGGGCATCGTGTCCGCCCAG 1200

Qy    269 CTGCCCAGGAGGAGTGGTGGCAGAGGAGGACCAGGACCCGTCGGAACTGAATCCC 1308
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CTGCCCAGGAGGAGTGGTGGCAGAGGAGGACCAGGACCCGTCGGAACTGAATCCC 1260

Qy    329 CAGAGAAAGCCAGGATCCTGCGCCCTTCCTGAACCGACTAGTTCGGCCTCGCAGA 1368
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CAGAGAAAGCCAGGATCCTGCGCCCTTCCTGAACCGACTAGTTCGGCCTCGCAGA 1320

Qy    389 CACCTAAAGGCCGGAAAAACACGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTT 1373
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CACCTAAAGGCGGAAAAACACGGGCTCGAAGAGCGATCGCAGCGCCATTATGAAGTT 1325

Qy    449 CACGACCTGGACAGGACGGAGCGCAGGCAGTGTGACGGACAGTGAGTGGCTGG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CACGACCTGGACAGGACGGAGCGCAGGCAGTGTGACGGACAGTGAGTGGCTGG

Qy    509 AAGCCAGAATCAAACAGCTCCAGCCCTCTGCGCTACACAACCGCCAGATCGGGGAGTTT
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AAGCCAGAATCAAACAGCTCCAGCCCTCTGCGCTACAACCGCCAGATCGGGGAGTTT

Qy    569 TCACCCGGCTGGGCTCTACTACTGTCAGGTGCACTTTGATGAGGGGAAG
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TCACCCGGCTGGGCTCTACTACTGTCAGGTGCACTTTGATGAGGGGAAG

Qy    629 TCTACCTGAAGCTGGACTTGCTGGTGGATGGTGTCTGGCCCTGGAG
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TCTACCTGAAGCTGGACTTGCTGGTGGATGGTGTCTGGCCCTGGAG

Qy    689 TCTCAGCCACTGCGGCCAGTTCCCTGGGCCCCAGCTCCGCCTCTGCCAGGTGTCT
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TCTCAGCCACTGCGGCCAGTTCCCTGGGCCCCAGCTCCGCCTCTGCCAGGTGTCT

Qy    749 TGTTGGCCCTGCGGGCTCTCCCTGCGGATCCGCACCCTCCCCTGGGCCCAT
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TGTTGGCCCTGCGGGCTCTCCCTGCGGATCCGCACCCTCCCCTGGGCCCAT

Qy    809 AGGCTGCCCCTTCCTCACCTACTTCGGACTCTTCGGACTCTTCACTGAGGGGCCCTG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 AGGCTGCCCCTTCCTCACCTACTTCGGACTCTTCGGACTCTTCACTGAGGGGCCCTG

Qy    869 CCCCACAGTCGTCCCGGCTGCCGGCGTGCCGGCTCTCCCTGCACAGCTCTGGGCACCCGGTC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CCCCACAGTCGTCCCGGCTGCCGGCGTGCCGGCTCTCCCTGCACAGCTCTGGGCACCCGGTC

Qy    929 CTGCCCCACCCTCAGCCGCGTCTTTGCTCCAGACCTGCCCCTCCCTTAGAGGCTGC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTGCCCCACCCTCAGCCGCGTCTTTGCTCCAGACCTGCCCCTCCCTCTAGAGGCTGC

Qy    989 GCCTGTTCACGTGTTTTCCATCCCACATAAATACAGTATTCCCACTCTTATCTTAC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GCCTGTTCACGTGTTTTCCATCCCACATAAATACAGTATTCCCACTCTTATCTTAC

Qy   1049 CCCCCACCGCCCACTCCCACCTCACTAGCTCCCAATCCCTGACCCCTTTGAGGCC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CCCCCACCGCCCACTCCCACCTCACTAGCTCCCAATCCCTGACCCCTTTGAGGCC

Qy   1109 GTGATCTCGACTCCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTC
```

CTAGCTCCCAATCCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCCCCCCTG 1128
|||||||||||||||||||||||||||||||||||||||||||||||||||||
CTAGCTCCCAATCCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCCCCCCTG 1080

CAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGTCCAGAAG 1188
|||||||||||||||||||||||||||||||||||||||||||||||||||||
CAGACCCCCAGGTCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGTCCAGAAG 1140

CACTTCAGGCACTAAGAGGGGCTGGACCTGCGGCCAGGAAGCCAAAGAGACTGGGC 1248
|||||||||||||||||||||||||||||||||||||||||||||||||||||
CACTTCAGGCACTAAGAGGGGCTGGACCTGCGGCCAGGAAGCCAAAGAGACTGGGC 1200

GCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAA 1308
|||||||||||||||||||||||||||||||||||||||||||||||||||||
GCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAA 1260

CTGTGTGATTTTTAAAACAGATATTATTTTTATTATTATTGTGACAAAATGTTGATA 1368
|||||||||||||||||||||||||||||||||||||||||||||||||||||
CTGTGTGATTTTTAAAACAGATATTATTTTTATTATTATTGTGACAAAATGTTGATA 1320

G 1373
G 1325

dard; DNA; 1306 BP.

(first entry)

e encoding tumour necrosis factor superfamily member.

tumour necrosis factor; osteopathic; bone; gene; ds.

2.

2002WO-US033022.

2001US-0329393P.

3-JEWISH HOSPITAL.

PF, Teitelbaum SL;

146/40.

ge 66-67; 78pp; English.

equence is that of a polynucleotide encoding a non-RANKL
tumour necrosis factor (TNF) superfamily. The invention
naturally-occurring proteins that contain one or more of the
ce loops of RANKL (see ABR42066-70) in combination with a
protein core obtained from a non-RANKL member of the TNF
also provided are polynucleotides encoding such proteins.
bind to RANK, acting as mimics of RANKL. They can be used to
formation by either inhibiting bone resorption or inducing
thus providing treatment for diseases or conditions
by loss of bone mass

c comprising a core, and at least one external loop, useful
processes of bone formation or inhibiting bone resorption,
f treatments for disease or condition characterized by loss

BP; 247 A; 434 C; 368 G; 257 T; 0 U; 0 Other;

XX
SQ   Sequence 1306 BP; 247 A; 434 C; 368 G; 257 T; 0 U; 0 Other;

Query Match   93.6%; Score 1285; DB 7; Length 1306;
Best Local Similarity 100.0%; Pred. No. 1e-263;
Matches 1285; Conservative   0; Mismatches   0; Indels   0;

```
Qy   89 CACAGCCCCCGCCCCCATGCCGCGCCGTCGGAGCCGGAGGGCGCGCCG
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 CACAGCCCCCGCCCCCATGCCGCGCCGTCGGAGCCGGAGGGCGGGCGCCG

Qy  149 AGCCGGCACCGCCCCTGCTGCTCCGCGCTGGGCCTGGCGCTGGCGCTGGC
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AGCCGGCACCGCCCCTGCTGCTCCGCGCTGGGCCTGGCGCTGGCGCTGGC

Qy  209 TCGGCTCCTGCTGCCGGTGTGCTCAGTTGGGGAGCCGGGCATGCGTCCGCCCA
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 TCGGCTCCTGCTGCCGGTGTGCTCAGTTGGGGAGCCGGGCATGCGTCCGCCCA

Qy  269 CTGCCCAGGAGGAGCTGGTGGCAGAGGAGGACCAGGACCCGTCGGAACTGAATCC
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CTGCCCAGGAGGAGCTGGTGGCAGAGGAGGACCAGGACCCGTCGGAACTGAATCC

Qy  329 CAGAGAAAGCCAGGATCCTGCGCCTTTCCTGAACCGACTAGTTCGGCCTCGCAG
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 CAGAGAAAGCCAGGATCCTGCGCCTTTCCTGAACCGACTAGTTCGGCCTCGCAG

Qy  389 CACCTAAAGGCCGGAAAAACACGGGCTGAAGAGGGATGCAGCCCATTATGAAGT
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CACCTAAAGGCCGGAAAAACACGGGCTGAAGAGGGATCGCAGCCCATTATGAAGT

Qy  449 CACGACCTGGACACGGAGCGCAGGCAGGTGTGGACGGGACAGTGAGTGGCTG
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 CACGACCTGGACACGGAGCGCAGGCAGGTGTGGACGGGACAGTGAGTGGCTG

Qy  509 AAGCCAGAGAATCAACAGCTCCAGCCCTCGCTCACAACCGCTACAACCGGGGAGTT
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 AAGCCAGAGAATCAACAGCTCCAGCCCTCGCTCACAACCGCCAGATCGGGGAGTT

Qy  569 TCACCCGGCTGGGCTTCTACTACCTGTACTGTCAGGTGCACTTTGATGAGGGGGAA
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 TCACCCGGCTGGGCTTCTACTACCTGTACTGTCAGGTGCACTTTGATGAGGGGGAA

Qy  629 TCTACCTGGAAGCTGACTTGCTGGTGGATGGTGTGCTGGCCCTCGCTGGA
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 TCTACCTGGAAGCTGACTTGCTGGTGGATGGTGTGCTGGCCCTCGCTGGA

Qy  689 TCTCAGCCACTGCGCCAGTCCCTCGGGCCCCAGTCTCCGCTCTGCCAGGTGTC
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 TCTCAGCCACTGCGCCAGTCCCTCGGGCCCCAGTCTCCGCTCTGCCAGGTGTC

Qy  749 TGTTGGCCCTGCGCGCCAGGGTCCTCCCTGCGGATCCGACCCGCCCCTGGGCCCA
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 TGTTGGCCCTGCGCGCCAGGGTCCTCCCTGCGGATCCGCACCCTCCCCTGGGCCCA

Qy  809 AGGCTGCGCCCCTTCCTCACCTACTTCGGACTCTTCCAGGTTCACTGAGGGCCCT
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 AGGCTGCGCCCCTTCCTCACCTACTTCGGACTCTTCCAGGTTCACTGAGGGGCCCT

Qy  869 CCCCACAGTCGTCCCAGGCTGCCGCGCCGGCTCTCCCGACACAGTCTCTGGGCACCCGGT
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 CCCCACAGTCGTCCCAGGCTGCCGCGCCGGCTCTCCCGACAGTCTCTGGGCACCCGGT

Qy  929 CTGCCCACCCTCAGCCGCTCTTTGCTCCAGACCTGCCCCTCCCTCCTCTAGAGGCTGC
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 CTGCCCACCCTCAGCCGCTCTTTGCTCCAGACCTGCCCCTCCCTCCTCTAGAGGCTGC

Qy  989 GCCTGTTCACGTGTTTCATCCCACATAAATACAGTATTCCCACTCTTATCTTA
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 GCCTGTTCACGTGTTTCATCCCACCATAAATACAGTATTCCCACTCTTATCTTA

Qy 1049 CCCCACCGGCCCCACTCCTCCACCTCACTAGTCCCCAATCCCTGACCCTTTGAGCC
        ||||||||||||||||||||||||||||||||||||||||||||||||
```

CC   antibody that binds specifically to endokine alpha polypeptide.
CC   present sequence represents a cDNA encoding a tumour necrosis fac
CC   family ligand.
XX
SQ   Sequence 1306 BP; 247 A; 434 C; 368 G; 257 T; 0 U; 0 Other;

Query Match   93.6%; Score 1285; DB 9; Length 1306;
Best Local Similarity 100.0%; Pred. No. 1e-263;
Matches 1285; Conservative 0; Mismatches 0; Indels 0; C

Qy   89 CACAGCCCCCGCCCCCATGGCCGCCGTCGGAGCCAGAGGGCGGAGGGGGCGCCGGC
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CACAGCCCCCGCCCCCATGGCCGCCGTCGGAGCCAGAGGGCGGAGGGGGCGCCGGC

Qy   149 AGCCGGGCACCGCCCTGCTGGTCCCGCTGCCGCTGGGCCTGGGCCTGGGCGCGCTGGCC
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AGCCGGGCACCGCCCTGCTGGTCCCGCTGCCGCTGGGCCTGGGCCTGGGCGCGCTGGCC

Qy   209 TCGGCCTTCCTGCCTGGTCAGTTTGGGGAGCCGGGCATCGTCTTCCGCCCAG
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TCGGCCTTCCTGCCTGGTCAGTTTGGGGAGCCGGGCATCGTCTTCCGCCCAG

Qy   269 CTGCCCAGGAGGAGCTGGTGGCAGAGGAGGAGCCAGGACCCGGTCGGAACTGAATCCC
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CTGCCCAGGAGGAGCTGGTGGCAGAGGAGGAGCCAGGACCCGGTCGGAACTGAATCCC

Qy   329 CAGAAGAAAGCCAGGATCCTCGGCCTTTCCTGAACCGACTAGTTCGGCCTCCGCAGA
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CAGAAGAAAGCCAGGATCCTCGGCCTTTCCTGAACCGACTAGTTCGGCCTCCGCAGA

Qy   389 CACCTAAAGGCCGGAAAACACGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTT
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CACCTAAAGGCCGGAAAACACGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTT

Qy   449 CACGACCTGGACAGGACGAGCGCAGCGAGGCAGGTGTGGACGGGACAGTGAGTGGCTGG
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CACGACCTGGACAGGACGAGCGCAGCGAGGCAGGTGTGGACGGGACAGTGAGTGGCTGG

Qy   509 AAGCCAGAATCAACAGCTCAGCCCCTGCGCTACACACGGCCAGATCGGGGAGTTT
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AAGCCAGAATCAACAGCTCAGCCCCTGCGCTACACACGGCCAGATCGGGGAGTTT

Qy   569 TCACCCGGCTGGGCTTCTACTACCTGTCAGGTGCACTTTGATGAGGGGAAG
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TCACCCGGCTGGGCTTCTACTACCTGTCAGGTGCACTTTGATGAGGGGAAG

Qy   629 TCTACCTGAAGCTGGACTTGCTGGTGGATGGTGTCTGGGCCCTGCCTGGAG
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TCTACCTGAAGCTGGACTTGCTGGTGGATGGTGTCTGGGCCCTGCCTGGAG

Qy   689 TCTCAGCCACTGCGGCCAGTTCCCTGGGCCCCAGCTCCGGCTCTGCCAGGTGTCT
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TCTCAGCCACTGCGGCCAGTTCCCTGGGCCCCAGCTCCGGCTCTGCCAGGTGTCT

Qy   749 TGTTGGCCCTGCGGCCAGGGTCCTCCCGGGATCCGCACCCTCCCCTGGGCCCAT
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TGTTGGCCCTGCGGCCAGGGTCCTCCCGGGATCCGCACCCTCCCCTGGGCCCAT

Qy   809 AGGCTGCCCCCTTCCACCTACTTCGGACTCTTCCAGGTTCACTGAGGGGCCCTG
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 AGGCTGCCCCCTTCCACCTACTTCGGACTCTTCCAGGTTCACTGAGGGGCCCTG

Qy   869 CCCCACAGTCGTCCCAGGCGTGCCGGCTCCCCTGACAGCTCTCTGGGCACCCGGTC
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 CCCCACAGTCGTCCCAGGCGTGCCGGCTCCCCTGACAGCTCTCTGGGCACCCGGTC

Qy   929 CTGCCCCACCCTCAGCCGCGCTCTTTGCTCCAGACCTGCCCCTCCCCTCCTTAGAGGCTGCC
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CTGCCCCACCCTCAGCCGCGCTCTTTGCTCCAGACCTGCCCCTCCCCTCCTTAGAGGCTGCC

Qy   989 GCCTGTTCACGTGTTTTCCATCCCACATAAATACAGTATTCCCACTCTTATCTTAC
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 GCCTGTTCACGTGTTTTCCATCCCACATAAATACAGTATTCCCACTCTTATCTTAC

Db   ... CACCGCCCACTCTTCCACCTCACTAGCTCCCAATCCCTGACCCTTTGAGGCCCCCA 1020

Qy   ... TCTGACTCCCCCTGGCCACAGACCCCCAGGGCATTGTTCACTGTACTCTGTG 1168
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   ... TCTGACTCCCCTGGCCACAGACCCCCAGGGCATTGTTCACTGTACTCTGTG 1080

Qy   ... AGGATGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGCA 1228
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db   ... AGGATGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGGGGCA 1140

Qy   ... GCCAAGAGACAGCTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAG 1288
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   ... GCCAAGAGACAGCTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAG 1200

Qy   ... GCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTATTTTATTATT 1348
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   ... GCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTATTTTTATTATT 1260

Qy   ... TGACAAAATGTTGATAAATGG 1373
             |||||||||||||||||||||
Db   ... TGACAAAATGTTGATAAATGG 1285

dard; cDNA; 1306 BP.

(first entry)

coding TNF ligand family member #12.

an; tumour necrosis factor; TNF ligand; endokine alpha;
e resorption disorder; osteoporosis; Paget's disease;
ification.

41.

C A.
LLI B.

Rosen CA,   Nardelli B;

072/66.
206.

EQ ID NO 23; 145pp; English.

relates to an isolated nucleic acid molecule encoding a
s factor family ligand. A composition comprising the
body or its fragment is used for treating an individual in
ased level of endokine alpha activity. The endokine alpha
esent in a heterotrimeric complex is used for treating an
ing a disorder associated with excessive bone resorption,
osis, Paget's disease or arterial calcification. Treating an
ring a disorder associated with insufficient bone resorption
nistering an endokine alpha antagonist, which is the

alpha gene useful for preparing a composition for treating a
ated with excessive or insufficient bone resorption e.g.,
Paget's disease or arterial calcification.

2002US-00218547.

2001US-0312542P.
2001US-0330761P.

affinity isolation of the corresponding receptor, the nucleic ac
which can be used to transform tumour cells to render them more
responsive to TREPA and to screen for TREPA mimics. Ribozymes,
RNA, antibodies or peptides, are used to treat TREPA-associate
diseases, e.g. tumours and metastases (by inhibiting vascularisa
inflammation or a wide range of autoimmune conditions, conditio
involving abnormal stimulation of epithelial cells (e.g.
atherosclerosis, for birth control (inhibiting ovulation and pl
formation) or other angiogenic conditions (e.g. ulcers)

```
CC
CC
CC
CC
CC
CC
CC
XX
SQ   Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 U; 0 Other;

Query Match             89.3%;  Score 1226.4;  DB 2;  Length 1236;
Best Local Similarity   99.5%;  Pred. No. 2.9e-251;
Matches 1230; Conservative    0; Mismatches    6;  Indels    0;

Qy   106 ATGGCCGCCCGTCGAGCCAGAGAGGCGGAGGGGCGCCGGGGGGAGCCGGCGCACCG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGGCCGCCCGTCGAGCCAGAGAGGCGGAGGGGCGCCGGGGGGAGCCGGCGCACCG

Qy   166 CTGGTCCGCGTCGCGCCTGGGCGCTGGGCCTGGCCTCGCGCTGGCCTCCGGCCTCTGC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 CTGGTCCGCGTCGCGCCTGGGCGCTGGGCCTGGCCTCGCGCTGGCCTCCGGCCTCCTGC

Qy   226 GTGGTCAGTTGGGGAGCCGGGCATCGCTGTTCCGCCCAGGAGCCTGCCCAGGAGG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTGGTCAGTTGGGGAGCCGGGCATCGCTGTTCCGCCCAGGAGCCTGCCCAGGAGG

Qy   286 GTGGCAGAGGAGGACCAGGACCCGTCGGAACTGAATCCCCAGACAGAAGAAAGCC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GTGGCAGAGGAGGACCAGGACCCGTCGGAACTGAATCCCCAGACAGAAGAAAGCC

Qy   346 CCTGCGCCTTTCCTGAACCGACTAGTTCGGCTCCGCAGAAGTGCACCTAAAGGCC
         |||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db   241 CCTGCGCCTTTCCTGAACCGACTAGTTCGGCTCCGAAGAAGTGCACCTAAAGGCC

Qy   406 ACACGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGAC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ACACGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGAC

Qy   466 GGAGCGCAGGCAGGTGTGGACGAGTAGTGCTGGGAGGAAGCCAGAATCA
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GGAGCGCAGGCAGGTGTGGACGGGACAGTAGTGCTGGGAGGAAGCCAGAATCA

Qy   526 TCCAGCCCTGCGCTACAACCGCCAGATCGGGAGTTATAGTCACCGGGGCTG
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TCCAGCCCTGCGCTACAACCGCCAGATCGGGAGTTTATAGTCACCCGGGGCTG

Qy   586 TACTACCTGTACTGTCAGTGACCAGTGACAGTGAGTGGCTGGGAGGAAGCCTGTCTACCTGAAGC'
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TACTACCTGTACTGTCAGTGACCACTTTGATGAGGGAAGGCTGTCGGGCTGTTCGGGCCTACCTGAAGC'

Qy   646 TTGCTGGTGGATGGTGCTGGGCCTGCGCTGCCTGGAATTCTCAGCCACTG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TTGCTGGTGGATGGTGCTGGGCCTGCGCTGCCTGGAGAATTCTCAGCCACTG

Qy   706 AGTTCCCTGGGCCCCAGCTGCCGCTCCGCCTCGCCGCTTGCTGGGCGTGTTGGCCTGCC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 AGTTCCCTGGGCCCCAGCTGCCGCTCCGCCTCGCCAGTGTCTGCCCAGGCGTGTTGGCCTACCTGAAGC'

Qy   766 GGGTCCTCCCTGCGGATCCGCACCCCATCTCAAGGCTGCCCCCTT
         |||||||||||||||||||||||||||||||||||||||||||
Db   661 GGGTCCTCCCTGCGGATCCGCACCTCCCCTGGCCCATCTCAAGGCGCCCCCTT

Qy   826 ACCTACTTCGGACTCTTCCAGGTCACTGAGGGGCCTTGGTCTCCCACAGTCGT(
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 ACCTACTTCGGACTCTTCCAGTTCACTGAGGGGCCCTTGGTCTCCCCGCAGTCGT(

Qy   886 GCTGCCGGCTCCCCTCCGACAGCTCTGGACACCGGCTCCCCCTGCCCCCACCCT(
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GCTGCCGGCTCCCCTCGACAGCTCTGGACACCCGGTCCCCTCTGCCCCCACCCT(

Qy   946 GCTCTTTGCTCCAGACCTGCCCCTCCCCTTCCAGACCTGGGCCTGTCTGCGCCTGTTCAGTC
```

CCACCGCCCACTCTCCACTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCCA 1108
|||||||||||||||||||||||||||||||||||||||||||||||||||
CCACCGCCCACTCTCCACTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCCA 1020

ATCTCGACTCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTG 1168
|||||||||||||||||||||||||||||||||||||||||||||||||||
ATCTCGACTCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTG 1080

AAGGATGGGTCAGAAGACCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGCGGCA 1228
|||||||||||||||||||||||||||||||||||||||||||||||||||
AAGGATGGGTCAGAAGACCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGCA 1140

AGCCAAAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAG 1288
|||||||||||||||||||||||||||||||||||||||||||||||||||
AGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAG 1200

AGCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTATTTTTATTATT 1348
|||||||||||||||||||||||||||||||||||||||||||||||||||
AGCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTATTTTTATTATT 1260

GTGACAAAATGTTGATAAAATGG 1373
||||||||||||||||||||||
GTGACAAAATGTTGATAAAATGG 1285

ndard; cDNA; 1236 BP.

(first entry)

endothelium proliferative agent gene.

othelium proliferative agent; TREPA; wound healing; cancer;
ing; vascularisation; apoptosis; autoimmune; birth control.

Location/Qualifiers
1..750
/*tag= a
/product= "TREPA"

98WO-US002859.

97US-00798692.
98US-00021706.

T LAB.

'255/38.
'745.

cleic acid encoding TREPA - useful for diagnosis and
autoimmune disease, tumours and inflammation.

ge 123-4; 142pp; English.

ed endothelium proliferative agent (TREPA), or its
agonists, are used to treat a deficit of TREPA, e.g. to
healing or tissue grafting, by promoting vascularisation,
e apoptosis for treating cancer and eliminating autoreactive
un adjunct to cancer chemotherapy or antiviral treatment.
s can also be used to target cytotoxic agents or for

CC particularly members of tumour necrosis factor (TNF) family mole
CC designated as TREPA (TNF related endothelium proliferative agent
CC Soluble biologically active TREPA are used to treat TREPA-associ
CC diseases, tumours or metastases. TREPA is used for inducing angi
CC in human for promoting wound healing and for vascularising graft
CC for successful grafting and to promote tissue grafts. The presen
XX sequence is a cDNA clone ID #690050 encoding human TREPA
SQ Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 U; 0 Other;

Query Match   89.3%; Score 1226.4; DB 4; Length 1236;
Best Local Similarity 99.5%; Pred. No. 2.9e-251;
Matches 1230; Conservative 0; Mismatches 6; Indels 0;

```
Qy  106 ATGGCGCCCGTCGGAGCCAGAGGCGGAGGGGGCCGGGGAGGGAGCCGGGCACGG(
            ||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db    1 ATGGCGCCCGTCGGAGCCAGAGCCGGAGGGGGCCGGGGAGGGAGCCGGGCACGG(

Qy  166 CTGGTCCCGCTCCGCTGGGCCTGGCCTGGCCTGGCCTGGCCTCGGCCTCTGCT
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTGGTCCCGCTCCGCTGGGCCTGGCCTGGCCTGGCCTGGCCTCGGCCTCTGCT

Qy  226 GTGGTCAGTTTGGGGAGCCGGGCATCGCTGTCCGCCCAGGAGCCTGCCCAGGAGG/
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GTGGTCAGTTTGGGGAGCCGGGCATCGCTGTCCGCCCAGGAGCCTGCCCAGGAGG/

Qy  286 GTGGCAGGAGGAGGACCAGGACCCGTCGGAACTGAATCCCAGACAGAAGAAAGCC/
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 GTGGCAGGAGGAGGACCAGGACCCGTCGGAACTGAATCCCAGACAGAAGAAAGCC/

Qy  346 CCTGCGCCTTTCCTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGCCG
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 CCTGCGCCTTTCCTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGCCG

Qy  406 ACACGGCTCGAAGAGCGATCGCAGCCATTATGAGAGTTCATCCAGACCTGGACA
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 ACACGGCTCGAAGAGCGATCCGCAGCCATTATGAGAGTTCATCCAGACCTGGACA

Qy  466 GGAGCGCAGGCAGGTGGACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAA
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GGAGCGCAGGCAGGTGGACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAA

Qy  526 TCCAGCCCTCTGCGCTACAACCGACCAGATCGGGGAGTTTATAGTCACCGGGCTGG
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 TCCAGCCCTCTGCGCTACAACCGACCAGATCGGGGAGTTTATAGTCACCCGGGCTGG

Qy  586 TACTACCTGTACTGTCAGGTGCACTTTGATGAGGGGAAGGCGTGTCTACCTGAAGCT
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 TACTACCTGTACTGTCAGGTGCACTTTGATGAGGGGAAGGCGTGTCTACCTGAAGCT

Qy  646 TTGCTGGTGGATGGTGTGCTGGCCCTGGCCCTGGCGCTGCTGGAGGAATTCAGCCACTGC
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 TTGCTGGTGGATGGTGTGCTGGCCCTGGCCCTGGCGCTGCTGGAGGAATTCAGCCACTGC

Qy  706 AGTTCCCTGGGGCCCCAGCTTCGCGCCTCTGCCAGGTGTCTGGGCTGTTGGCCCTGCG/
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 AGTTCCCTGGGGCCCCAGCTTCACTGAGGGGCCCTGGCCGTCTGGGCTGTTGGCCCTGCG/

Qy  766 GGGTCCTCCTGCGATCCGCACCCTCCCCTGGCCCATCTCAAGGTCCTCAAGGCTGCCCCTT(
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 GGGTCCTCCTGCGATCCGCACCCTCCCCTGGCCCATCTCAAGGTCCTCAAGGCTGCCCCTT(

Qy  826 ACCTACTTCGGACTCTTCCAGGTTCACTGAGGGGCCCTGGTTCTCCCCACCAGTGTCC
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 ACCTACTTCGGACTCTTCCAGGTTCACTGAGGGGCCCTGGTTCTCCCCGGCGTGTCG/

Qy  886 GCTGCCGGCTCCCCTGCACAGCTTCTCGGGCACCCGGTCCCCTGCCCACCCTC/
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 GCTGCCGGCTCCCCTGCACAGCTCTCTGGGCACCCGGTCCCCTGCCCACCCTC/

Qy  946 GCTCTTTGCTCCAGACCTGCCCCTCCCTCTAGAGGCTGCCTGGGCCTGTTCACGTGT
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 GCTCTTTGCTCCAGACCTGCCCCTCCCTCTAGAGGCTGCCTGGGCCTGTTCACGTGT
```

```
TTTGCTGCCAGACCTGCCCCTCCCTCTAGAGGCTGCCTGGGCCTGTTCACGTGTTT  900
CCCACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCACTCT  1065
CCCACATAAATACAGTATTCCCACTCTTATCTTACAACACCCCACCGCCACTCT   960
CTCACTAGCTCCCCAATCCTGACCCTTTGAGGCCCCCAGTGACTCTGACTCCCCC  1125
CTCACTAGCTCCCCAATCCTGACCCTTTGAGGCCCCCAGTGACTCTGACTCCCCC  1020
CCACAGACCCCCAGGGCATTGTGTTCACTGTACTGTACTCTGTGGGCAAGGATGGGTCCAG  1185
CCACAGACCCCCAGGGCCATTGTGTTCACTGTACTGTACTCTGTGGGCAAGGATGGGTCCAG  1080
CCCCACTTCAGGCACTAAGAGAGGGCTGACCTGGCCGGCAGGAAGCCAAAGAGACTG  1245
CCCCACTTCAGGCACTAAGAGAGGGCTGACCTGGCCGCGCAGGAAGCCAAAGAGACTG  1140
TAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGA  1305
TAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGA  1200
TCCCTGTGTGGATTTTTAAAACAGATATTATTTT  1341
TCCCTGTGTGGATTTTTAAAACAGATATTATTTT  1236
```

dard: cDNA; 1236 BP.

(first entry)

TNF related endothelium proliferative agent) cDNA.

necrosis factor; TNF; angiogenesis; wound healing; TREPA;
ndothelium proliferative agent; tumour; metastasis;
nerary; ss.

Location/Qualifiers
1..750
/*tag= a
/product= "Human TREPA (TNF related endothelium
proliferative agent)"

98US-00105343.

97US-00798692.
98US-00021706.

LAB.

60/29.
91.

genesis in mammal at desired sites for promoting wound
ministering soluble fragment of extracellular domain of
factor related endothelium proliferative agent protein.

73-74; 53pp; English.

vention relates to extracellular signal molecules,

their active fragments. APO4 is useful for diagnosing prostate c
determining levels of APO4 in an individual. Prostate cancer can
treated using APO4 selective binding agents linked to a therapeu
moiety. APO4 polypeptides are also useful for identifying select
binding agents, useful in diagnosis/treatment of disease by bind
agents to the polypeptide/active fragment which is extracellular
expressed on the cell surface. The binding is preferably perform
vivo. APO4 polypeptides/ active fragments are also useful for sc
for agonists and antagonists by binding and observing the change
activity. Effective pharmacological agents useful in diagnosis o
treatment of disease are also identified using APO4 polypeptides
fragments and APO4 signal transducer molecules that specifically
with a cytoplasmic domain of APO4 and detecting a change in leve
activity. The method is performed in vivo or in vitro. APO polyp
are all useful as immunogens for preparing antibodies. APO4 is a
useful for diagnosis/treatment of developmental or gestational
abnormalities. APO8 was transfected to human breast carcinoma ce
MCF-7, and induced apoptosis

XX
SQ   Sequence 1030 BP; 223 A; 317 C; 279 G; 211 T; 0 U; 0 Other;

Query Match   60.7%;   Score 833.4;   DB 2;   Length 1030;
Best Local Similarity   99.9%;   Pred. No. 1.1e-167;
Matches 834;   Conservative   0;   Mismatches   1;   Indels   0;

```
Qy    229  GTCAGTTTGGGGAGCCGGGCATCGCTGTCCGCCCAGGAGCCTGCCCAGGAGGAGC
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  GTCAGTTTGGGGAGCCGGGCATCGCTGTCCGCCCAGGAGCCTGCCCAGGAGGAGC

Qy    289  GCAGAGGAGGACCAGGACCCGTGGAACTGAATCCCAGACAGAGAAAGCCAGG
                ||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db     61  GCAGAGGAGGACCAGGACCCGTGGAACTGAATCCCAGACAGAAGAAAGCCAGG

Qy    349  GCGCCTTCCTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGA
                ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  GCGCCTTCCTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGA

Qy    409  CGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCAGACGACCTGGACAGG
                |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  CGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCAGACCTGGACAGG

Qy    469  GCGCAGGCAGGTGTGGACGGACGAGTGAGTGGCTGGGAGGAAGCCAGAATCAACA
                ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  GCGCAGGCAGGTGTGGACGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACA

Qy    529  AGCCCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGCTGGGC
                ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  AGCCCTGCGCTACTACCGCCAGATCGGGGAGTTTATAGTCACCCGGCTGGGC

Qy    589  TACCTGTACTGTCAGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGG
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  TACCTGTACTGTCAGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGG

Qy    649  CTGGTGGATGGTGTGCTGCCCTGCGCTCGCCTGGAGAATTCTCAGCCACTGGGC
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  CTGGTGGATGGTGTGCTGCCCTGCGCTCGCCTGGAGAATTCTCAGCCACTGGGC

Qy    709  TCCCTCGGGCCCCAGCTCCGCCTCGCCCTCGCCATCCCCGGGCCATCCAAGGCTGCCCGGCGCC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  TCCCTCGGGCCCCAGCTCCGCCTCGCCCTCGCCATCCCCGGGCCATCCAAGGCTGCCCGGCGCC

Qy    769  TCCTCCCTGCGGATCCGACTCCCCGCCGCACCCCTGGGCCCATCTCAAGGCTGCCCCCTTCC
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  TCCTCCCTGCGGATCCGACTCCCCGCCGCACCCCTGGGCCCATCTCAAGGCTGCCCCCTTCC

Qy    829  TACTTCGGACTCTTCCAGTTCACTGAGGGCCCTGGTCTCCCCACAGTCGTCCA
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  TACTTCGGACTCTTCCAGTTCACTGAGGGGCCCTGGTCTCCCCACAGTCGTCCA

Qy    889  GCCGGCTCCCCTCGACAGCTCTCTGGGCACCCGGTCCCCTCTGCCCACCCTCAGC
                ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  GCCGGCTCCCCTCGACAGCTCTCTGGGCACCCGGTCCCCTCTGCCCACCCTCAGC
```

(left column, rotated sequence continuations)

```
CCCCACATAAATACAGTATTCCCACTCTTATCTTACAAACTCCCCACCGCCCACTCT    1065
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCCACATAAATACAGTATTCCCACTCTTATCTTACAAACACCCACCGCCCACTCT      960

CCTCACTAGCTCCCCAATCCCTTGAGGCCCCAGTGATCTCGACTCCCCC            1125
      ||||||||||||||||||||||||||||||||||||||||||||||||||
CCTCACTAGCTCCCCAATCCCTGACCCCTTGAGGCCCCAGTGATCTCGACTCCCCC     1020

GCCACAGACCCCCAGGCATTGTGTTCACTGTACTTCTGTGGGCAAGGATGGGTCCAG     1185
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCCACAGACCCCCAGGGCATTGTGTTCACTGTACTTCTGTGGGCAAGGATGGGTCCAG    1080

ACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGCGCAGGAAGCCAAAGAGACTG    1245
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
ACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGCGCAGGAAGCCAAAGAGACTG    1140

CTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCCCTTGA      1305
      |||||||||||||||||||||||||||||||||||||||||||||||||||||
CTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCCCTTGA      1200

CTCCCTGTGGATTTTTAAAAACAGATATTATTTT                          1341
      ||||||||||||||||||||||||||||||||
CTCCCTGTGGATTTTTAAAACAGATATTATTTT                           1236
```

...dard; DNA; 1030 BP.

(first entry)

DNA.

...sis factor receptor; signal transducer molecule; TNF; APO4;
...abnormality; gestational abnormality; prostate cancer;
APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
...domain; immunogen; antibody preparation; breast carcinoma;
...man; ss.

Location/Qualifiers
1..627
/*tag= a
/product= "TNRL3"

98WO-US018393.

97US-00924634.

WASHINGTON.

...191/17.
...590.

...rosis Factor family receptor polypeptides and ligands -
...agnosis and treatment of prostate cancer and developmental
...l abnormalities.

Fig 13A; 156pp; English.

...n describes isolated Tumor Necrosis Factor (TNF) family
...peptides: APO4, APO6, APO8 and APO9 or their active
...d isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or

```
Qy  106  ATGGCCGCCCGTCGGAGCCAGAGCCGGAGGGGGCCGGGGGAGCCGGGCACCG(
              ||||||||||||||||||||||||||||| |||||||||||||| ||||||||
Db    1  ATGGCCGCCCGTCGGAGCCAGAGCCGGAGGGGGCGGGGGAGCGCGGGCACCG(

Qy  166  CTGGTCCCGCTGCGCTGGGCGTGGGGCCTGGGCCTGGGCCTGGGCCTCCGGC(
              |||||||||||| ||||||||||||| ||||||||||||||||||||||||||
Db   61  CTGGCCCCGCTGCGTGCTGAGCCTGGGCCTGGGCCTGGGCCTGGGCCTCCTGC(

Qy  226  GTGGTCAGTTTGGGGAGCCGGGCATCGTGTCCGCCCAGGAGCCTGCCCAGGAGG/
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db  121  GTGGTCAGCCTGGGGACGCTGGGCAACCGTGTCTGCCCAGGAGCCTTCTCAGGAGG/

Qy  286  GTGGCAGAGGAGGACCAGGACCCGTCGGAACTGAATCCCAGACAGAAGAAAGCC/
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  ACAGCAGAGGAGGACCGCCGGGGAGCCCCCTGAACTGAATCCCAGACAGAGGAAAGCC/

Qy  346  CCTGCGCCTTCCTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCC(
              ||||||||||||||||||||||||||||| |||||||| ||||||||||||||
Db  241  GTGGTACCTTTCTTGGAACAACTAGTCCGGCCTCGCAGAAGTGCTCCTAAAGGCCC(

Qy  406  ACACGGGCTCGAAGAGGCGATCGCAGCCCATTATGAAGTTCATCCAGCACCTGGACA
              | ||||| ||||||||||||||||| ||||||||||||||||||||||||||||||
Db  301  GCGCGGCCTTCGCCGAGCTATTGCAGCCCATTATGAGGTTCATCCTCGGCCAGGACA

Qy  466  GGAGCGCAGGCAGGTGTGGACGGACAGTGAGTGGTCTGGGAGGAAGCCAGAATCAA
              ||| |||||| |||||||||| ||||||||||| |||||||||| ||||||||||
Db  361  GGAGCACAAGCAGGTGTGGATGGGACAGTGAGTGGTCTGGGAAGAGACCAAAATCAA

Qy  526  TCCAGCCCTCTGCGCTACACCGCCAGATCGGGAGTTTATAGTCACCCGGCTGG
              |||||||||||||||||||||||||||||| |||||| |||||||||||||||
Db  421  TCCAGCCCTCTGCGCTACGCCGCCAGATTGGGAATTTACAGTCATCAGGGCTGG

Qy  586  TACTACCTGTACTGTCAGGTGCACTTTGATGAGGGGAAGGCTGTCACCTGAAGCT
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  TACTACCTGTACTGTCAGGTGCACTTTGATGAGGGGAAAGGCTGTCACCTGAAGCT

Qy  646  TTGCTGGTGGATGGTGTGCTGGCCCTGCGCGTCCTGGAGGAATTCTCAGCCACTGC
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db  541  TTGCTGGTGAACGGTGTGCTGGCCCTGCGCGTCCTGGAAGAATTCTCAGCCACAGC

Qy  706  AGTTCCCTCGGGCCCCAGCTCCGCCTCGCCCAGGTGTCTGGGGTTGGGCCCTGCG
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601  AGTCTTCCCTCGGGCCCCAGCTCGCCTCGTTGTGCCCAGGTGTCTGGGCGTGTGCCGCTGCG

Qy  766  GGGTCCTCCGCTGCGGATCGGCAGCATCACCACCTCCCATCCAAGGCTGCCCCTT
              ||||||| ||||||||||||||| ||||||||| ||||||||||||||||||||
Db  661  GGGTCTTCCCTTGGATCGGCAGCATCACCACCTCATCTTAAGGCTGCCCCCTT

Qy  826  ACCTACTTCGGACTCTTCAGGTTCACTGAGGGGCCCTGGTCTCCCCACAGTCGTC
              |||||||||||| ||||||| ||||||||||||||||| ||||| |||| ||||
Db  721  ACCTACTTTGGACTCTTTCAGTTCAAGTTCACTGAGGGGCTCTTCACTGCTCTCCCCAGATTCCTTA

Qy  886  GCTGCCGGCTCC-----CCTCGACAGCTCTCTGGGCACCCGGTCCCCTGCCCCA(
              |||||||||||       ||||||||||||||||||||||||||||||||||||
Db  781  TTCCCTGGCTCCCAGGAGCATCACCACCTCCCTACCCCACACCCCCACTCCTCCACC(

Qy  941  CAGCCGCTCTTTGTTGCTCCAGACCTGCCCCCCTCCCTCTAGAGGCTGCCTGGGCCTGTTC/
              || ||| ||||||||||| |||||||||||||||| |||||||||||||||||||||||
Db  841  C-GCTGCTCCTTGGTCCAGTCCTGTCTCCC--TCAAAGGCAGCCAGAGCTTGTTC/

Qy 1001  GTTTTCCATCCCACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACC(
              |||| |||||                |||||||||| ||||||||||||||||||
Db  898  GTTTCCATTCC------ACAGACGTATCCTTGCTCTTCTTAACATCCCATCCCACC(

Qy 1061  ACTCTCCACCTACTAGCTCCCAATCCTGACCCTTTGAGGCCCCCAGTGATCTCC
              |||||||||||||||||||||||||||||||           --TTATCCCTC
Db  951  ACTATCCACCTCACTAGCTCCCAAAGCCCCTAC----------TTATCCCTC

Qy 1121  CCCCCCTGGCCACAGACCCCCAGGGCATTGTTTCACTGTACTCTGTGGGCAAGGA(
              |||||                                 ||||||||||||||||
Db  996  CCCCACCCACTTCACCCGACCCACGTGTTTTATTGACTTTGTGCAC---------

Qy 1181  TCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGGCGGAGGAAGCCA/
```

Left-column text fragments:

GCTCCAGACCTGCCCCTCCCTTCAGAGGCTGCCTGGGCCTGTTCACGTGTTTCCA 1008
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCTCCAGACCTGCCCCTCCCTTCAGAGGCTGCCTGGGCCTGTTCACGTGTTTTTCCA 780

ACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCCCACT 1063
||||||||||||||||| ||||||||||||||| |||||||||||||||||
ACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCCACT 835

dard; cDNA; 1239 BP.

(first entry)

AK coding sequence.

EAK; TNF relatedness and weak ability to induce cell death;
ecrosis Factor; TWEAK; fibrosis; cardiac disease;
; lung disease; kidney disease; skin disease;
nal tract disease; pancreatic disease;
organ disease; neural disease; cartilage disease;
connective tissue disease; cellular death; hepatotropic;
l; gastrointestinal; osteopathic; gene; ss.

Location/Qualifiers
1..750
/*tag= a
/product= "Fl-TWEAK"

A2.

2003WO-US011350.

2002US-0371611P.

I INC.

ubowski A, Zheng T, Hahm K;

56/78.

712.

EAK-related condition, e.g. liver, gastrointestinal, kidney,
ic, cartilage or neural tissue condition in a subject
nistering to the subject a TWEAK agonist or antagonist.

D ID NO 2; 120pp; English.

quence is the coding sequence for murine transmembrane FL-
atedness and weak ability to induce cell death, where TNF
osis Factor). TWEAK is a member of the TNF family. TWEAK
tagonists are useful for treating a TWEAK-related
. fibrosis; cardiac disease; liver disease; lung disease;
; skin disease; skeletal muscle disease; adipose tissue
ointestinal tract disease; pancreatic disease; reproductive
neural disease; cartilage disease; bone disease;
sue disease; cellular death; and a pathological condition
pressing a TWEAK receptor.

BP; 249 A; 386 C; 331 G; 273 T; 0 U; 0 Other;
rity   49.8%;   Score 683.4;   DB 9;   Length 1239;
       77.0%;   Pred. No. 9.7e-136;
nservative   0; Mismatches 221; Indels 70; Gaps 9;

CC   leucine zipper multimerisation domain, and the extracellular dom
CC   human TWEAK. The fusion protein was used in the isolation of hum
CC   receptor (TWEAKR)-expressing clones from a COS cell human cDNA l
CC   The TWEAK protein is a member of the tumour necrosis factor (TNF
CC   and induces angiogenesis. TWEAKR may therefore be used to screen
CC   develop TWEAK agonists and antagonists for the modulation of
CC   angiogenesis, to be used in the treatment and diagnosis of human
CC   The disorders mediated by angiogenesis include ocular disorders
CC   characterised by ocular neovascularisation such as diabetic reti
CC   neovascular glaucoma, retinoblastoma, retinopathy of prematurity
CC   retrolental fibroplasia, rubeosis, uveitis, macular degeneration
CC   corneal graft neovascularisation, and inflammatory diseases such
CC   arthritis, rheumatism and psoriasis. Other treatable diseases in
CC   malignant and metastatic conditions such as sarcomas and carcino
CC   benign tumours and preneoplastic conditions, myocardial angiogen
CC   haemophilic joints, scleroderma, vascular adhesions, atheroscler
CC   plaque neovascularisation, telangiectasia, wound granulation, co
CC   atherosclerosis, peripheral atherosclerosis and ischaemia
XX
SQ   Sequence 898 BP; 187 A; 266 C; 267 G; 178 T; 0 U; 0 Other;

Query Match   45.8%;   Score 629.2;   DB 4;   Length 898;
Best Local Similarity 99.5%;   Pred. No. 3.1e-124;
Matches 631;   Conservative   0; Mismatches   3; Indels   0;

Qy   232 AGTTTGGGAGCCGGGCATCGTCTTCCGCCCAGGAGCCTGCCCAGGAGGAGCTGG
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   250 AGTTTGGGAGCCGGGCATCGTGTTCCGCCCAGGAGCCTGCCCAGGAGGAGCTGG

Qy   292 GAGGAGGACCAGGACCCGTCGGAACTGAATCCCCAGACAGAAGAAAGCCAGGATC
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   310 GAGGAGGACCAGGACCCGTCGGAACTGAATCCCCAGACAGAAGAAAGCCAGGATC

Qy   352 CCTTTCCTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAA
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   370 CCTTTCCTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGGAAAA

Qy   412 GCTCGAAGAGCGATCGCAGCCATTATGAAGTTCATCCACGACCTGGACAGGACG
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   430 GCTCGAAGAGCGATCGCAGCCATTATGAAGTTCATCCACGACCTGGACAGGACG

Qy   472 CAGGCAGGTGTGACGGGACAGTGAGTGGTGGAGGAAGCCAGAATCAACAGCT
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   490 CAGGCAGGTGTGACGGGACAGTGAGTGGTGGAGGAAGCCAGAATCAACAGCT

Qy   532 CCTCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGCTGGGCTCT
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   550 CCTCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGCTGGGCTCT

Qy   592 CTGTACTGTCAGGTGCACTTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGGACT
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   610 CTGTACTGTCAGGTGCACTTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGGACT

Qy   652 GTGGATGGTGTGCTGGCCCTGGCCTGCTGGGAGGAATTCTCAGCCACTGCGGCCAC
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   670 GTGGATGGTGTGCTGGCCCTGGCCTGCTGGGAGGAATTCTCAGCCACTGCGGCCAC

Qy   712 CTCGGGCCCCAGCTCCGCCTCTGCCAGGTGTCTGGGCTGTTGGCCCTGCGGCCAGG
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   730 CTCGGGCCCCAGCTCCGCCTCTGCCAGGTGTCTGGGCTGTTGGCCCTGCGGCCAGG

Qy   772 TCCCTGCGGATCCGCACCCTCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCAC
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   790 TCCCTGCGGATCCGCACCCTCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCAC

Qy   832 TTCGGACTCTTCCAGGTTCACTGAGGGGCCCTGG 865
            ||||||||||||||||||||||||||||||
Db   850 TTCGGACTCTTCCAGGTTCACTGAGCGGCCGCGG 883

RESULT 14
AAV18599
ID AAV18599 standard; cDNA; 1168 BP.

                    ||||||| ||| |||||||||| |||||||||||||||||| |||
             --------CAGGCACTGAGATGGGCTGGACCTGGTGGCAGGAAGCCAGAGA 1082
TGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCCGAGA-AACAAGACAAGCTCCTC 1299
||| ||||||||| ||||||| |||||||||||||| ||||||||||||| |||
TGGGACTAGCCAGAAGTTCCCAACTGTGAGGGGGAGAGCTGGGGACCAAGCTCCTC 1142

TGAGAATTCCCTGTGGATTTTAAAACAGATATTATTTTTATTATTATTGTGACAAA 1359
||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
GGA---TCCCTGTGGATTTGAAA--AGATACTATTTTATTATTATTGTGACAAA 1196

TTGA 1366
||||
TTAA 1203

ndard; DNA; 898 BP.

(first entry)

ector pDC409-LZ-TWEAK fusion protein-encoding DNA.

ellular domain; tumour necrosis factor; TNF; angiogenesis;
scularisation; diabetic retinopathy; neovascular glaucoma;
na; retinopathy of prematurity; retrolental fibroplasia;
itis; macular degeneration; arthritis; rheumatism; ds;
neovascularisation; psoriasis; metastatic condition;
nour; sarcoma; carcinoma; benign tumour; haemophilic joint;
condition; myocardial angiogenesis; wound granulation;
vascular adhesion; telangiectasia; ischaemia; human;
ic plaque neovascularisation; coronary atherosclerosis;
heroclerosis; pDC409-LZ-TWEAK; TWEAK receptor; TWEAKR;
n.

Location/Qualifiers
52..873
/*tag= a
/product= "Fusion protein comprising a growth hormone
leader, a leucine zipper multimerisation domain, and
human TWEAK extracellular domain"

2000WO-US034755.

99US-0172878P.
2000US-0203347P.

IEX CORP.

975/44.
499.

giogenesis in a mammal for treating diseases mediated by
e.g. solid tumors and vascular deficiencies of cardiac or
ssue, by administering antagonist or agonist of TWEAK

ge 39-40; 46pp; English.

represents a DNA from the expression vector pDC409-LZ-TWEAK,
a fusion protein comprising a growth hormone leader, a

(first entry)

tumour necrosis factor related ligand (TRELL) gene.

necrosis factor related ligand; tnf; treatment; cancer;
sease; immune system; stimulation; suppression;
on; ds.

Location/Qualifiers
2..679
/*tag= a
/note= "tumour necrosis factor related ligand"

97WO-US013945.

96US-0023541P.
96US-0028515P.
97US-0040820P.

N INC.
GENEVA FACULTY MEDICINE.

e Y, Browning JL;

619/13.
524.

is factor related ligand - useful for, e.g. treating cancer,
isease and immune responses to tissue grafts.

45-46; 69pp; English.

is that encoding mouse tumour necrosis factor related ligand
l or active fragments can be included with a carrier in
l compositions to treat cancer, autoimmune diseases or
ses to tissue grafts, or to stimulate or suppress the immune
useful to screen for TRELL receptors, by labelling with a
el and screening compositions for binding. Agents
ith TRELL-receptor binding can also be screened for, can
istered, optionally with interferon- gamma, to induce cell
;, suppress or alter immune responses (especially involving
rcinoma cells) involving a signal pathway between TRELL and
The DNA sequence can be used in gene therapy for TRELL-
ders in mammals (especially humans), e.g. tumours,
d inflammatory diseases or inherited genetic disorders, by
nto cells, and expressing, therapeutically effective amounts
.g. a virus comprising a gene encoding TRELL. It may also
the preparation of prepare probes for screening
atic DNAs for TRELL-encoding sequences and for antisense

BP; 242 A; 360 C; 298 G; 268 T; 0 U; 0 Other;

arity   44.8%;   Score 614.6; DB 2;   Length 1168;
nservative   75.8%;   Pred. No. 4.2e-121;
0;   Mismatches 219;   Indels   70;   Gaps   9;

Qy   298 GACCAGAGACCCCTGCGGAACTGAATCCCAGACAGAGAAGAAGCCAGGATCCTGCGC(
           ||||||||||||||||||||||||||||||||||||
Db   122 CGCGGGGAGCCCCCTGAACTGAATCCCCAGACAGAGGAAAGCCAGGATGGGTAC(

Qy   358 CTGAAACCGACTAGTTCCGGCCTGCAGAGAGTGCACCTAAAGGCCGAAAACACGGG(
           ||| ||| ||||||| ||||| |||||| |||||||| |||||||
Db   182 TTGGAACAACTAGTCCGGCCTGCAGAGAGTGCTCCTAAAGGCCGAAGGCGCGGGCC(

Qy   418 AGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACGGAGCGC/
           ||||||||||| |||||| |||||| ||||||||||||||||||||||||||||
Db   242 CGAGCTATTGCAGCCCATTATGAAGTTCATCCTCGGCCAGGACAGCAGGAGCAC/

Qy   478 GGTGTGACGGACAGTGAGTGGCTGGGAGAAGCCAGAATCAACAGCTTCCAGCCC
           |||||||||||| |||||||||||||||||||| |||||| ||||||||||||
Db   302 GGTGTGGATGGGACAGTGAGTGGCTGGGAAGAACCAAAATCAACAGCTTCCAGCCC

Qy   538 CGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGCTGGGCTTCTACTACCT
           ||||| |||||||||||||| |||| |||| ||||| ||| |||||| ||||||||
Db   362 CGCTACGACCGCCAGATTGGGGAATTTACAGTCATCAGGCTCATCAGGGCTCTACTACCT

Qy   598 TGTCAGGTGCACTTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTGGT
           |||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db   422 TGTCAGGTGCACTTTGATGAGGGGAAAGGCTGTCTACCTTGAAGCTGGACTTGCTGGT

Qy   658 GGTGTGTGGCCCTGCGCTGCCTGCCTGGGAGGAATTCTCAGCCACTGCGCGCCAGTTCCCT
           ||||||||||||||| ||| |||| |||||| ||||||||||||||||||||||||
Db   482 GGTGTGTGGCCCTGCGCTGCCTGCCTGGGAGAATTCTCAGCCACAGCACGCAAGCTCTCC

Qy   718 CCCCAGTCCCTGCCTCTGCCAGTGTCTGTCCAGTGTTGGCCTGTTGGCCTGCCAGGGTCCTC
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   542 CCCCAGCTCCGTTTGTGCCAGTGTGTCTGGGCTGTTGCCCGTGTGGCGGCCAGGGTCTTC

Qy   778 CGGATCCGCACCCTCCCCCTGGGCCCCATCTCAAGGCTGCCCCCCTTCCTCACCTACTT
           |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   602 CGGATCCGCACCCTCCCCCTGGGCTCATCTTAAGGCTGCCCCCCTTCCTAACCTACTT

Qy   838 CTCTTCCAGGTTCACTGAGGGGGCCCCTGGTCTCCCACAGTCGTCCCAGGCTGCCGG
           ||||||||||||||||||||||||||| ||||||||||| |||||||||
Db   662 CTCTTTCAAGTCACTGAGGGGCCTTGCTCTCCCAGATTCCTTAAACTTTCCCTGG

Qy   898 -----CCTCGACAGCTCTCGTGGGCACCCGGTCCCCCTGCCCCTCAGCCGCT
                |||||| ||| ||||||||||||||||||||||||||||||||
Db   722 AGGAGCATCACCACACCTCCCTACCCCCACCCCACTCCTCACCCCCTC-GCTGCT

Qy   953 GCTCCAGACCTGCCCCTCCCCTCTAGAGGCTGCCTGGGCCTGTTCACGTGTTTCCA
           ||||||||||||||||||||||||||||| ||||||||||||||||||
Db   781 GGTCCAGTCCTGCTCTCTCC--TCAAAGGCAGCCAGGAGCTTGTTCACATGTTCCAT

Qy   1013 ACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCCCACTCTCCA
           |||| ||||||||||| ||| ||| ||| || ||||| |||||||||||
Db   838 ------ACAGACGTATCCTTGCTCTTCTTAACATCCCATCCACCAACTATCCA

Qy   1073 ACTAGTCCCCAATCCTGACCCTTTGAGGCCCCAGTGATCTCGACTCCCCCTG
           |||||||| |||||||| ||||||||||||||||||||| ||
Db   892 ACTAGCTCCCAAAGCCCTAC-----------------TTATCCCTGACTCCCCCACC

Qy   1133 CAGACCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAG.
           |||||| |||||||||||||||||||||||||||||||
Db   937 CACCCGACCACGTGTTTATTGACTTTGTGCAC------------------

Qy   1193 CACTTTCAGGCACTAAGAGAGGGGCTGACCTGCGGCAGGAAGCCAAAGAGACTGGGC(
           ||||||| ||||||||||||||||||||||| |||||
Db   969 -----CAGGCACTGAGATGGCTGACCTTGGTGCACTGGTGGCAGGAGAACCTGGGA(

Qy   1253 GCCAGGAGTTCCCAAATGTGAGGGGCGAGA-AACAAGACAAGCTCTCCCTTGAGA/
           |||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   1024 GCCAGAAGTTCCCAACTGTGAGGGGGAAGAGCTGGGAGCAAGCTCTCCCTGGA---

Qy   1312 CCTGTGGATTTTAAAACAGATATTATTTTTATTATTGTGACAAAATGTTGA
           ||||||||||||||| ||||||||||||||||||||||||||||||||
Db   1080 CCTGTGGATTTTGAAA--AGATACTATTTTTATTATTGTGACAAAATGTTAA

ndard; DNA; 701 BP.

(first entry)

DNA.

sis factor receptor; signal transducer molecule; TNF; APO4;
1 abnormality; gestational abnormality; prostate cancer;
APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
domain; immunogen; antibody preparation; breast carcinoma;
ouse; ss.

Location/Qualifiers
1..636
/*tag= a
/product= "TNRL3"

98WO-US018393.
97US-00924634.
WASHINGTON.

3191/17.
3591.

rosis Factor family receptor polypeptides and ligands -
agnosis and treatment of prostate cancer and developmental
l abnormalities.

Fig 13B; 156pp; English.

n describes isolated Tumor Necrosis Factor (TNF) family
peptides: APO4, APO6, APO8 and APO9 or their active
d isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
fragments. APO4 is useful for diagnosing prostate cancer by
evels of APO4 in an individual. prostate cancer can also be
; APO4 selective binding agents linked to a therapeutic
polypeptides are also useful for identifying selective
s, useful in diagnosis/treatment of disease by binding of
 polypeptide/active fragment which is extracellular, or
the cell surface. The binding is preferably performed in
lypeptides/ active fragments are also useful for screening
and antagonists by binding and observing the changer in APO4
ective pharmacological agents useful in diagnosis or
disease are also identified using APO4 polypeptides/active
APO4 signal transducer molecules that specifically interact
asmic domain of APO4 and detecting a change in level of APO4
method is performed in vivo or in vitro. APO polypeptides
l as immunogens for preparing antibodies. APO4 is also
agnosis/treatment of developmental or gestational
. APO8 was transfected to human breast carcinoma cell line
duced apoptosis

BP; 139 A; 210 C; 203 G; 149 T; 0 U; 0 Other;

37.8%;   Score 519.2;   DB 2;   Length 701;
arity 87.3%;   Pred. No. 7.3e-101;
onservative   0;   Mismatches 83;   Indels   0;   Gaps   0;

```
Qy   220  CTGGCCGTGGTCAGTTTGGGGAGCCGGCATCGCTGTCCGCCCAGGAGCCTGCCC
               ||| |||||||||||  ||||||| || |||| ||||| ||||||||||||| |
Db     1  CTGGTCGTGGTCAGCCTGGGGAGCTGGGCAACGCTGTCTGCCCAGGAGCCTTCTC

Qy   280  GAGCTGGTGGCAGAGGAGGACCAGGACCCGTTCGGAACTGAATCCCCAGACAGAAG
               ||||||  || ||||||||  ||  ||||| |||||||||||||||||||||||||
Db    61  GAGCTGACAGCAGAGGACCGCCGCGGGAGGCCCCCTGAACTGAATCCCCAGACAGAGG

Qy   340  CAGGATCCTCGCGCCTTCCTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACCTA
               ||||||  | | |||||| |||  | ||||| | |||||||| ||||||| ||||
Db   121  CAGGATGTGGTACCTTTCTTGGACAACTAGTCCGGCCTCGAAGAAGTGCTCCTA

Qy   400  CGGAAAACACGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGAC
               |||||  | ||||  ||| ||| || ||||| |||||||||||| ||||||||||
Db   181  CGGAAGGGCGGGCCTCGCCAGCTATTGCAGCCCATTATGAGGTTCATCCTCGGC

Qy   460  CAGGACGGAGCGCAGGCAGGTGTGGACGGGACAGTGAGTCGGCTGGGAGGAAGCCA
               ||||| ||||| ||| ||||||||| | |||||||||||| |||||||| || |||
Db   241  CAGGATGGAGCACAAGCAGGGTGTGATGGACAGTGAGTGCTGGGAAGAGACCA

Qy   520  AACAGCTCCAGCCCCTGCGGCTACAACCGGCCAGATCGGGGAGTTTATAGTCACCC
               ||||||||||||| |||||||||| |||||||||| | ||||| |||| |||| ||
Db   301  AACAGCTCCAGCCCTCTGCGGCTACGACCGGCCAGATTGGGGAATTTACAGTCATCA

Qy   580  GGGCTCTACTACCTGTACTGTCAGTGCACTTTGATGAGGGAAGGCTGTCTACC
               |||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db   361  GGGCTCTACTACCTGTACTGTCAGTGCACTTTGATGAGGGAAAGCTGTCTACC

Qy   640  CTGGACTTGCTGGTGGATGTGTCTGGCCCTGGCCCTGGCGGAATTCTCAG
               ||||||||||||||  | ||| |||||||| ||| ||||||||| |||||||
Db   421  CTGGACTTGCTGGTGAACCGTGTGCTGGGCCCTGCGCTGCTGCGGAAGAATTCTCAG

Qy   700  GCGGCCAGTTCCCTCGGGCCCCAGCTCCGCTCTGCCAGGTGTCGGGGCTGTTGG
               || |||||| |||| ||||||||||||||  | |||||||||||| ||| |||||
Db   481  GCACCAAGCTCTCCTGGGCCCCAGCTCCGTTGTGCCAGGTGTCCAGGTGTCTTGG

Qy   760  CGGCCAGGGTCCTCCCTGCGGATCCGCACCCTCCCCTGGGCCATCTCAAGGCTG
               |||||||||||| ||||| |||||||||||||||||||||||| ||||| |||||
Db   541  CGGCCAGGGTCTTCCCTTCGGATCCGCACCCTCCCCTGGGCTCATCTTAAGGCTG

Qy   820  TTCCTCACCTACTTCGGACTCTTCCAGGTTCACTGAGGGGCCCTGGTCTCCC  87
               ||||| |||||||| ||||||||| ||||||||||||||||||||| |||||||
Db   601  TTCCTAACCTACTTTGGACTCTTTCAAGTTCACTGAGGGGCCCTTGCTCTCCC  65
```

Search completed: April 7, 2004, 21:32:21
Job time : 528.669 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

...ic search, using sw model

...ril 7, 2004, 23:15:35 ; Search time 523.048 Seconds
                          (without alignments)
                          9837.847 Million cell updates/sec

-09-245-198A-3
73
atgtcattgtagactttga........gacaaaatgttgataaatgg 1373

ENTITY_NUC
...pop 10.0 , Gapext 1.0

70632 seqs, 1873875610 residues

...ts satisfying chosen parameters: 4941264

...gth: 0
...gth: 2000000000

...inimum Match 0%
...aximum Match 100%
...isting first 45 summaries

...ublished Applications NA:*
/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
/cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
/cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
/cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
/cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
/cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
/cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
/cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
/cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
/cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
/cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
/cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
/cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
/cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
/cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
/cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

...the number of results predicted by chance to have a
...r than or equal to the score of the result being printed,
...ed by analysis of the total score distribution.

SUMMARIES

...ery ...ch	Length	DB	ID	Description
3.0	1364	9	US-09-822-849A-19	Sequence 19, Appl
3.2	1353	14	US-10-210-951-3	Sequence 3, Appli
3.2	1353	14	US-10-211-884-3	Sequence 3, Appli
3.6	1306	12	US-10-202-062-23	Sequence 23, Appl
3.6	1306	14	US-10-272-411-16	Sequence 16, Appl
3.6	1306	14	US-10-218-547-23	Sequence 23, Appl
3.6	1306	14	US-10-272-328A-16	Sequence 16, Appl
3.6	1306	14	US-10-310-793-29	Sequence 29, Appl
3.8	898	9	US-09-742-454A-1	Sequence 1, Appli
3.8	898	9	US-09-883-777-1	Sequence 1, Appli
3.4	493	10	US-09-918-995-21225	Sequence 21225, A
3.5	408	9	US-09-960-352-2197	Sequence 2197, Ap
3.7	213	12	US-10-085-783A-55176	Sequence 55176, A
3.7	213	15	US-10-242-535A-55176	Sequence 55176, A
3.7	264	9	US-09-983-965-2183	Sequence 2183, Ap

C	16	67.2	4.9	3163	14	US-10-017-161-1857	Sequence
C	17	67.2	4.9	3163	15	US-10-292-798-1513	Sequence
C	18	65.4	4.8	1064	9	US-09-804-682-29	Sequence
C	19	63.8	4.6	5452	14	US-10-017-161-1481	Sequence
C	20	63.8	4.6	5452	15	US-10-292-798-1189	Sequence
C	21	63.4	4.5	778	12	US-10-424-599-54839	Sequence
C	22	62.2	4.5	12733	14	US-10-032-393-47	Sequence
C	23	62.2	4.5	12739	14	US-10-032-393-8	Sequence
C	24	62	4.5	815	12	US-10-424-599-20495	Sequence
C	25	61.4	4.5	1117	14	US-10-017-161-1403	Sequence
C	26	61.4	4.5	1117	15	US-10-292-798-1141	Sequence
C	27	60	4.4	60	10	US-09-908-975-13797	Sequence
C	28	60	4.4	60	10	US-09-908-975-31339	Sequence
C	29	60	4.4	60	10	US-09-908-975-31341	Sequence
C	30	60	4.4	60	10	US-09-908-975-31596	Sequence
C	31	60	4.4	60	10	US-09-908-975-31597	Sequence
C	32	59.6	4.3	3133	14	US-10-017-161-1483	Sequence
C	33	59.6	4.3	3133	15	US-10-292-798-1191	Sequence
C	34	56	4.1	1065	9	US-09-804-682-33	Sequence
C	35	55.6	4.0	771	12	US-10-424-599-107958	Sequence
C	36	55.2	4.0	580	12	US-10-424-599-2320	Sequence
C	37	55	4.0	791	12	US-10-424-599-64474	Sequence
C	38	54.4	4.0	1327	15	US-10-398-221-1265	Sequence
C	39	54	3.9	931	14	US-10-198-846-8585	Sequence
C	40	53.2	3.9	53522	10	US-09-904-968A-1	Sequence
C	41	52.6	3.8	636	12	US-10-424-599-72906	Sequence
C	42	52.2	3.8	594	14	US-10-142-426-10	Sequence
C	43	52.2	3.8	594	14	US-10-123-155-10	Sequence
C	44	52.2	3.8	594	14	US-10-146-731-10	Sequence
C	45	52.2	3.8	594	14	US-10-140-472-10	Sequence

ALIGNMENTS

RESULT 1
US-09-822-849A-19
; Sequence 19, Application US/09822849A
; Patent No. US20020045170A1
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakar
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEIN
; FILE REFERENCE: GIN 6403
; CURRENT APPLICATION NUMBER: US/09/822,849A
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/195,582
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 598
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 1364
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-849A-19

Query Match   98.0%; Score 1345.8; DB 9; Length 1364;
Best Local Similarity   99.9%; Pred. No. 0;
Matches 1347; Conservative 0; Mismatches 2; Indels 0; G:

Qy   25 TCCGCCCGCGGCTCCCCCCCCCCCCGATCCCTCGGGTCCCGGGATGGGGGGCCGGGGT(
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5 TCCGCCCGCGGCTCCCCCCCTCCCCCCGATCCCTCGGGTCCCGGGATGGGGGGCCGGGT(

Qy   85 CAGGACAGCACAGCCCCCGCCCCGCCCCCCATGGCCGCGCCGTCGGAGCCAGAGGCGGAGGGCGGG(
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Qy   1225  GGCAGGAAGCCAAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGCC
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1205  GGCAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGCC

Qy   1285  CAAGACAAGCTCCTCCCTTGAGAATTCCTGTGGATTTTAAAACAGATATTATT
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1265  CAAGACAAGCTCCTCCCTTGAGAATTCCTGTGGATTTTTAAAACAGATATTATT

Qy   1345  TATTATTGTGACAAAATGTTGATAAATGG 1373
           ||||||||||||||||||||||||||||
Db   1325  TATTATTGTGACAAAATGTTGATAAATGG 1353
```

RESULT 2
US-10-210-951-3
; Sequence 3, Application US/10210951
; Publication No. US20030170228A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi,Avi J.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Hillan,Kenneth J.
; APPLICANT: Marsters,Scot A.
; APPLICANT: Pan,James
; APPLICANT: Pitti,Robert M.
; APPLICANT: Roy,Margaret Ann
; APPLICANT: Smith,Victoria
; APPLICANT: Stone,Donna M.
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF
; FILE REFERENCE: P2931R1C1
; CURRENT APPLICATION NUMBER: US/10/210,951
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/014699
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 60/026943
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 60/059121
; PRIOR FILING DATE: 1997-07-17
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/062037
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; Remaining Prior Application data removed - See File Wrapper or PALM
; NUMBER OF SEQ ID NOS: 258
; SEQ ID NO 3
; LENGTH: 1353
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-210-951-3

  Query Match   96.2%; Score 1320.2; DB 14; Length 1353;
  Best Local Similarity   99.8%; Pred. No. 0;
  Matches 1322; Conservative   0; Mismatches   3; Indels   0; (

Qy   49 CGATCCCTCGGGTCCCGGGATGGGGGCGGGGTGAGGCAGGCAGGCAGCGCCCCCGCCCC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CGATCCCTCGGGTCCCGGGATGGGGGCGGGGTGAGGCAGGCAGGCAGCGCCCCCGCCCC

Qy   109 GCCGCCCCTCGGAGCCCAGAGGCGGGAGGGGCCGAGGGGGCGCCGGGGGAGCCGGCACCGCCCC
```

```
Qy   124  GCACAGCCCCCGCCCCATGGCCGCCCGTCGGAGGCGGAGGGGCCCCGG
          |||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy   204  GAGCCGGGCACCGCCGCCTCGTCGTCCGCTCGCCTGGGCCTGGGCGCTGGGCC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   184  GAGCCGGGCACCGCCGCCTCGTCGTCCGCTCGCCTGGGCCTGGGCGCTGGGCC
```

```
Qy   264  CTCGGCCTCTGCTGCCGTGGTCAGTTTGGGGAGCCGGGCATCGCTGTCGGCCAG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   244  CTCGGCCTCTGCTGGCCGTGGTCAGTTTGGGGAGCCGGGCATCGCTGTCCGGCCCAG
```

```
Qy   324  CCTGCCCAGGAGGAGAGCTGGTGGCAGAGGAGGACCAGGACCCGTCGGGAACTGAATCCC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   304  CCTGCCCAGGAGGAGAGCTGGTGGCAGAGGAGGACCAGGACCCGTCGGGAACTGAATCCC
```

```
Qy   384  ACAGAGAAAGCCAGGATCCTGCGCCTTTCCTGAACCGACTAGTTCGGCCTCGCAGA
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   364  ACAGAGAAGAAGCCAGGATCCTGCGCCTTTCCTGAACCGACTAGTTCGGCCTCGCAGA
```

```
Qy   444  GCACCTAAAGGCCGGAAAACACGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTT
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   424  GCACCTAAAGGCCGGAAAACACGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTT
```

```
Qy   504  CCACGACCTGGACAGGACGGAGCGCAGGCGCAGTGTGACGCAGTAGTGAGTGGCTGG
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   484  CCACGACCTGGACAGGACGGAGCGCAGGCGCAGTGTGACGCAGTAGTGAGTGGCTGG
```

```
Qy   564  GAAGCCAGAATCAACAGTCCAGCCCTCTGCGCTACAACCGCCAGATCGGGGAGTTT
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   544  GAAGCCAGAATCAACAGTCCAGCCCTCTGCGCTACAACCGCCAGATCGGGGAGTTT
```

```
Qy   624  GTCACCCGGGCTGGGCTTCTACTACCTGTACTGTCAGTGCACTTTGATGAGGGGAAG
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   604  GTCACCCGGGCTGGGCTTCTACTACCTGTACTGTCAGTGCACTTTGATGAGGGGAAG
```

```
Qy   684  GTCTACCTGAAGCTGGACTTGCTGGTCGATGGTGTCGCCCTGGCCTGCCTGGAG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   664  GTCTACCTGAAGCTGGACTTGCTGGTCGATGGTGTCGCCCTGGCCTGCCTGGAG
```

```
Qy   744  TTCTCAGCCACTGCGGCCAGTTCCCTCGGGCCCCAGCTCCGCCTCTGCCAGTGTCT
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   724  TTCTCAGCCACTGCGGCGAGTTCCCTCGGGCCCCAGCTCCGCCTCTGCCAGTGTCT
```

```
Qy   804  CTGTTGGCCCTGCGGCCAGGGTCCTCCCTGCGGATCCGACACCCTCCCCTGGGCCCAT
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   784  CTGTTGGCCCTGCGGCCAGGGTCCTCCCTGCGGATCCGACACCCTCCCCTGGGCCCAT
```

```
Qy   864  AAGGCTGCGCCCCCTTCCTCACCTACTTCGGACTCTTCCAGGTTCACTGAGGGCCCTG
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   844  AAGGCTGCGCCCCCTTCCTCACCTACTTCGGACTCTTCCAGGTTCACTGAGGGCCCTG
```

```
Qy   924  CCCCACAGTGTCGCGGCCAGTTGCGCGGCTCCCCTGACAGCTCTCTGGGCACCCGGTC
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   904  CCCCACAGTGTCGCGGCCAGTTGCGCGGCTCCCCTGACAGCTCTCTGGGCACCCGGTC
```

```
Qy   984  CCTGCCCCACCCTCAGCCGCTCTTGCTCCAGACCTGCGCCCCTCCCCTCCCCTAGAGGCTGC
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   964  CCTGCCCCACCCTCAGCCGCTCTTGCTCCAGACCTGCGCCCCTCCCCTCCCCTAGAGGCTGC
```

```
Qy  1044  GGCCTGTTCACGTGTTTTCCATCCCACATAAATACAGTATTCCCACTCTTATCTTAC
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1024  GGCCTGTTCACGTGTTTTCCATCCCACATAAATACAGTATTCCCACTCTTATCTTAC
```

```
Qy  1104  TCCCCACCACGCCCACTCTCCACCTCACTAGTTCCCCAATCCTGACCCTTTGAGGCC
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1084  TCCCCACCACGCCCACTCTCCACCTCACTAGTTCCCCAATCCTGACCCTTTGAGCC
```

```
Qy  1164  AGTGATCTCGACTCCCCCCGTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTC
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1144  AGTGATCTCGACTCCCCCCGTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTC
```

```
Qy  1224  GGGCAAGGATGGGTCCAGAGACCCCACTTCAGGCACTAAGAGGGGCTGACCTGGC
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1204  GGGCAAGGATGGGTCCAGAGACCCCACTTCAGGCACTAAGAGGGGCTGACCTGGC
```

```
Db   1141 ACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGCAGGAAGCCAAAGAGAC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1249 CTAGGCCAGGAGTTCCCAAATGTGAGGGCGGAGAAACAAGACAAGCTCCTCCCTT
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTAGGCCAGGAGTTCCCAAATGTGAGGGCGGAGAAACAAGACAAGCTCCTCCCTT
Qy   1309 TTCCCTGTGGATTTTTAAAACAGATATTATTTATTATTGTGACAAAATGT
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 TTCCCTGTGGATTTTTAAAACAGATATTATTTATTATTGTGACAAAATGT
Qy   1369 AATGG 1373
          |||||
Db   1321 AATGG 1325
```

RESULT 3
US-10-211-884-3
; Sequence 3, Application US/10211884
; Publication No. US20030175900A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Marsters, Scot A.
; APPLICANT: Pan, James
; APPLICANT: Pitti, Robert M.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stone, Donna M.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF T
; FILE REFERENCE: P2931R1C1
; CURRENT APPLICATION NUMBER: US/10/211,884
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/014699
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 60/026943
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 60/059121
; PRIOR FILING DATE: 1997-07-17
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/062037
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; Remaining Prior Application data removed - See File Wrapper or PALM
; NUMBER OF SEQ ID NOS: 258
; SEQ ID NO 3
; LENGTH: 1353
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-211-884-3

  Query Match   96.2%; Score 1320.2; DB 14; Length 1353;
  Best Local Similarity   99.8%; Pred. No. 0;
  Matches 1322; Conservative 0; Mismatches 3; Indels 0; Ga

Qy   49 CGATCCCTCGGGTCCCGGGATGGGGGGGGGGGGGGGGGGGGGGGGGCACAGCCCCCCGCCC
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CGATCCCTCGGGTCCCGGGATGGGGGGGATGGGGGGCGGGTGAGGCAGGCACAGCCCCCCGCCC
```

```
          |||||||||||||||||||||||||||||||||||||||||||||||
CCCGTCCGAGCCAGAGAGCCGGAGAGGGGCGGCGGGGGGAGCCGGCCACCGCCCTGCTG 120
CGCTCGCGCTGGGCCTGGGCCGCTGGCCGCTGCTGCTGCTCCGCGCCTGCTGCGCCGTG 228
          |||||||||||||||||||||||||||||||||||||||||||||||||
CGCTCGCGCTGGGCCTGGGCCGCTGGCCGCTGCTGCTGCTCCGCGCCTGCTGCGCCGTG 180
GTTTGGGGAGCCGGGCATCGCTGTCTGCCCCAGGAGCCTGCCCAGGAGGAGTTGGTG 288
          ||||||||||||||||||||||||||||||||||||||||||||||||
GTTTGGGGAGCCGGGCATCGCTGTCTGCCCCAGGAGCCTGCCCAGGAGGAGTTGGTG 240
AGGAGGACCAGGACCCGTCGGAACTGAATCCCAGACAGAGAAAGCCAGGATCCT 348
          |||||||||||||||||||||||||||||||||||||||||||||||
AGGAGGACCAGGACCAGGTCTCGCGGAACTGAATCCCAGACAGAGAAGCCAGGATCCT 300
CTTTCCTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGAAAACA 408
          ||||||||||||||||||||||||||||||||||||||||||||||
CTTTCCTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGAAAACA 360
CTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGACAGACGGA 468
          ||||||||||||||||||||||||||||||||||||||||||||||
CTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGACAGACGGA 420
AGGCAGGTGTGACGGGACAGTGAGTGGCTGGGAGAAGCCAGAATCAACAGCTCC 528
          |||||||||||||||||||||||||||||||||||||||||||||
AGGCAGGTGTGACGGGACAGTGAGTGGCTGGGAGAAGCCAGAATCAACAGCTCC 480
CTCTGCGCTACAAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTAC 588
          |||||||||||||||||||||||||||||||||||||||||||||
CTCTGCGCTACAAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTAC 540
CGTACTGTCAGGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGACTTG 648
          ||||||||||||||||||||||||||||||||||||||||||||||
CGTACTGTCAGGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGACTTG 600
CGGATGGTGTGCTGGGCCCTGCGCTGCCTGCTGGAGGAATTCTCAGCCACTGCCGCCAGT 708
          ||||||||||||||||||||||||||||||||||||||||||||||||
CGGATGGTGTGCTGGGCCCTGCGCTGCCTGCTGGAGGAATTCTCAGCCACTGCCGCGAGT 660
CGGGGCCCCAGCTCGCCCTCGCCCTCTGCCAGGTGTCTGGGCGTGTTGGCCCTGGCCCAGGG 768
          |||||||||||||||||||||||||||||||||||||||||||||
CGGGGCCCCAGCTCGCCCTCTGCCCAGGTGTCTGGGCGTGTTGGCCCTGGCCCAGGG 720
CCTCGCGGATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACC 828
          ||||||||||||||||||||||||||||||||||||||||||||||
CCTCGCGGATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACC 780
CGGACTCTTCCAGTTCACTGAGGGCCCCTGGTCTGTCCCACAGTCGTCCCAGGCT 888
          ||||||||||||||||||||||||||||||||||||||||||||||
CGGACTCTTCCAGTTCACTGAGGGGCCCCTGGTCTGTCCCAGTCGTCCCAGGCT 840
CCTCCCCTGACAGCTCTCTGGGCACCCGGTCCCCCTCTGCCCACCCTCAGCCCGCT 948
          ||||||||||||||||||||||||||||||||||||||||||||||
CCTCCCCTGACAGCTCTCTGGGCACCCGGTCCCCCTCTGCCCACCCTCAGCCCGCT 900
CCTCCAGACCTGCCCCTCCCCTCCCTCTCCTAGAGGCTGCCTGGGCCTGTGTTCACGTGTTTTCCA 1008
          |||||||||||||||||||||||||||||||||||||||||||||
CCTCCAGACCTGCCCCTCCCCTCCCTCCTAGAGGCTGCCTGGGCCTGTTCACGTGTTTTCCA 960
CATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCCAGCCGCCCACTCTCCA 1068
          |||||||||||||||||||||||||||||||||||||||||||||
CATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCCAGCCGCCCACTCTCCA 1020
CTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCCCCCCTG 1128
          |||||||||||||||||||||||||||||||||||||||||||||
CTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCCCCCCCTG 1080
AGACCCCCAGGTCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTTCCAGAAG 1188
          |||||||||||||||||||||||||||||||||||||||||||||
AGACCCCCAGGTCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTTCCAGAAG 1140
ACTTCAGGCACTAAGAGGGGCTGGACCTGGCGCAGGAAGCCAAAGAGACTGGGC 1248
```

```
Db   1141 ACCCCACTTCAGGCACTAAGAGAGGGCTGGACCTGGCGGCAGGAAGCAAGAGAC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1249 CTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCTT
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCTT

Qy   1309 TTCCCTGTGGATTTTAAAACAGATATTATTTTATTATTATTGTGACAAAATGT
          ||||||||||||||| ||||||||||||||| |||||||||||||||||||||
Db   1261 TTCCCTGTGGATTTTTAAAACAGATATTATTTTTATTATTATTGTGACAAAATGT

Qy   1369 AATGG 1373
          |||||
Db   1321 AATGG 1325
```

RESULT 4
US-10-202-062-23
; Sequence 23, Application US/10202062
; Publication No. US20040038349A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.,
; TITLE OF INVENTION: Heteromultimeric TNF Ligand Family members
; FILE REFERENCE: PF559
; CURRENT APPLICATION NUMBER: US/10/202,062
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,838
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patent-in version 3.0
; SEQ ID NO 23
; LENGTH: 1306
; TYPE: DNA
; ORGANISM: human
US-10-202-062-23

Query Match   93.6%; Score 1285; DB 12; Length 1306;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1285; Conservative   0; Mismatches   0; Indels   0;

```
Qy    89 CACAGCCCCCGCCCCATGGCCCCGTCGGACCGCAGAGGCGGAGGGGCGGCCG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 CACAGCCCCCGCCCCATGGCCCCGTCGGACCGCAGAGGCGGAGGGGCGGCCG

Qy   149 AGCCGGGCACCGCCCCTGCTGGTCTCCGCGCCTCGGGCCTGGGCCTGGC
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 AGCCGGGCACCGCCCCTGCTGGTCTCCGCGCCTCGGGCCTGGGCCTGGC

Qy   209 TCGGCCTCGTCGGTGTCGCCGTCCGTGTCGGTCAGTGGGCCTGTCGCCCA
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TCGGCCTCGTCGGTGTCGCCGTCCGTGTCGGTCAGTGGGCCTGTCGCCCA

Qy   269 CTGCCCAGGAGGAGCTGGTGGCAGAGGAGGACCAGGACCCGTCGGAACTGAATCC
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CTGCCCAGGAGGAGCTGGTGGCAGAGGAGGACCAGGACCCGTCGGAACTGAATCC

Qy   329 CAGAAGAAAGCCGGAAAACACGGCCTCGCCTTCCTGAAGAGCTCGAAGAGCTTCGCGAG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CAGAAGAAAGCCGGAAAACACGGCCTCGCCTTCCTGAAGAGCTCGAAGAGCTTCGCGAG

Qy   389 CACCTAAAGGCCGGAAAACACGGCCCATTATGAAGTGCACCTAAAGGCCGGAAAACA
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CACCTAAAGGCCGGAAAACACGGCCCATTATGAAGTGCACCTAAAGGCCGGAAAACA

Qy   449 CACGACCTGGACAGGACCGGAGCCGGCAGGCAGTGTGGACGGATCGGAGTGAGTGGCTG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CACGACCTGGACAGGACCGGAGCCGGCAGGCAGTGTGGACGGATCGGAGTGAGTGGCTG

Qy   509 AAGCCAGATCAACAGCTTCCGCGCTCACCCGGCTGGCTCTCCAGGCAAGATCAACAGCTC
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AAGCCAGATCAACAGCTTCCGCGCTCACCCGGCTGGCTCTCCAGGCAAGATCAACAGCTC

Qy   569 TCACCCGGGCTGGGCTCTACTACCCTGTACTGGTGCACTTTGATGAGTGGGGAC...
```

```
Qy   109 CCCGTCGGAGCCGGAGAGGGGCGCGGGAGCCGGGGAGGGGCGCACCGCCCCTGCTG 168
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    65 CCCGTCGGAGCCGGAGAGGGGCGCGGGAGCCGGGGAGGGGCGCACCGCCCCTGCTG 120

Qy   169 CCGCTCGCGCTCGGGCCTGGGCCTGGCTGCTCGCGCCTCTCCTGCTGCCGGCCGTG 228
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CCGCTCGCGCTCGGGCCTGGGCCTGGCTGCTCGCGCCTCTCCTGCTGCCGGCCGTG 180

Qy   229 AGTTTGGGAGCCGGGCATCGCTGTCGCCCAGGAGCTGCCCAGGAGGAGCTGGTG 288
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AGTTTGGGAGCCGGGCATCGCTGTCGCCCAGGAGCTGCCCAGGAGGAGCTGGTG 240

Qy   289 GAGGAGACCAGGACCCGTCGGAACTGAATCCCACAGAAGAAAGCCGGAGATCCT 348
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GAGGAGACCAGGACCCGTCGGAACTGAATCCCACAGAAGAAAGCCGGAGATCCT 300

Qy   349 CCTTTCCTGAACCGACTAGTTCGGCCTCGCAGCAGTGCACCTAAAGGCCGGAAAACA 408
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CCTTTCCTGAACCGACTAGTTCGGCCTCGCAGCAGTGCACCTAAAGGCCGGAAAACA 360

Qy   409 GCTCGAAGAGCGATCGCAGCCCATTATGAAGTCATCCACAGACCTGGACAGGACCGA 468
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GCTCGAAGAGCGATCGCAGCCCATTATGAAGTCATCCACAGACCTGGACAGGACCGA 420

Qy   469 CAGGCAGGTGTGGACGGGACAGTGAGTGGCTGGTGGGACGTGAGCCAGAATCAACAGCTCC 528
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CAGGCAGGTGTGGACGGGACAGTGAGTGGCTGGTGGGACGTGAGCCAGAATCAACAGCTCC 480

Qy   529 CCTCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTAC 588
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CCTCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTAC 540

Qy   589 CTGTACTGTCAGGTGCACTTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGGACTTG 648
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CTGTACTGTCAGGTGCACTTTGATGAGGGGAAGGCTGTCTACCTGAAGCTGGACTTG 600

Qy   649 GTGGATGGTGTGTGCTGGCCCTGCCCGCTGCCCTGGAGGAATTCTCAGCCACTGCGGCCAGT 708
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GTGGATGGTGTGTGCTGGCCCTGCCCGCTGCCCTGGAGGAATTCTCAGCCACTGCGGCGAGT 660

Qy   709 CTCGGCGCCCCAGCTGCCTTCGCGCTCTCGCCTCAGGTGTCTGGGCTGTTGGCCTGCCCAGGG 768
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CTCGGGCGCCCCAGCTGCCTTCGCGCTCTCGCCTCAGGTGTCTGGGCTGTTGGCCTGCCCAGGG 720

Qy   769 CCCCTGCGGATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACC 828
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CCCCTGCGGATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACC 780

Qy   829 TTCGGACTCTTCCAGGTTCACTGAGGGGCCCTGGTCTCCCCACAGTCGTCCCAGGCT 888
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 TTCGGACTCTTCCAGGTTCACTGAGGGGCCCTGGTCTCCCCACAGTCGTCCCAGGCT 840

Qy   889 GGCTCCCCTGCACAGCTGTCTGGGCACCCGGTCCCCGTCTCCCTGCGCCCACCCTCAGCCGCT 948
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GGCTCCCCTGCACAGCTGTCTGGGCACCCGGTCCCCGTCTCCCCACCCTCTGCCCACCCTCAGCCGCT 900

Qy   949 CGGCTCCAGACCTGCCCCTCTAGAGGCTGCCTGGGCCTGGCCTGTTCACGTGTTTTCCA 1008
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 CGGCTCCAGACCTGCCCCTCTAGAGGCTGCCTGGGCCTGGCCTGTTCACGTGTTTTCCA 960

Qy  1009 ACACATAAATACAGTATTCCCACTCTTTATCTTTACAACTCCCCACCGCTCCTCCA 1068
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 ACACATAAATACAGTATTCCCACTCTTTATCTTTACAACTCCCCACCGCTCCTCCA 1020

Qy  1069 CACTAGCTCCCCAATCCTGACCCTTTGAGGCCCCAGTGATCTGACTCGCCCCCCTG 1128
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 CACTAGCTCCCCAATCCTGACCCTTTGAGGCCCCAGTGATCTGACTCGCCCCCCTG 1080

Qy  1129 ACAGACCCCCAGGCATTGTTGTTCACTGTACTCTGTGGGCAAGGATGGTCCAGAAG 1188
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 ACAGACCCCCAGGCATTGTTGTTCACTGTACTCTGTGGGCAAGGATGGTCCAGAAG 1140

Qy  1189 CACTTCAGGCACTAAGAGAGGGCTGGACCTGGCGGCAGGAAGCAAGAGAC 1248
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 TCACCCGGGCTGGGCTCTACTACCCTGTACTGGTGCACTTTGATGAGGGGAC 1248
```

; SEQ ID NO 16
; LENGTH: 1306
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NCBI/AF030099.1
; DATABASE ENTRY DATE: 1997-12-20
; RELEVANT RESIDUES: (1)..(1306)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NCBI/NM_003809.2
; DATABASE ENTRY DATE: 2002-10-07
; RELEVANT RESIDUES: (1)..(1306)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NCBI/AF055872.1
; DATABASE ENTRY DATE: 1998-05-04
; RELEVANT RESIDUES: (1)..(1306)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NCBI/BC019047.1
; DATABASE ENTRY DATE: 2001-12-11
; RELEVANT RESIDUES: (1)..(1306)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NCBI/AF030100.1
; DATABASE ENTRY DATE: 1997-12-20
; RELEVANT RESIDUES: (1)..(1306)
US-10-272-411-16

Query Match   93.6%;   Score 1285;   DB 14;   Length 1306;
Best Local Similarity   100.0%;   Pred. No. 0;   Indels   0;   C
Matches 1285;   Conservative   0;   Mismatches   0;

```
Qy    89  CACAGCCCCCGCCCCATGGCCGCCCGTCGGAGCCCAGAGGCGGAGGGGCGCGG
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  CACAGCCCCCGCCCCATGGCCGCCCGTCGGAGCCCAGAGGCGGAGGGGCGCGG

Qy   149  AGCCGGGCACCGCCCTGGTCCCCGGCTCGGGCCTGGCCTGGCGCGCTGGCC
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  AGCCGGGCACCGCCCTGGTCCCCGGCTCGGGCCTGGCCTGGCGCGCTGGCC

Qy   209  TCGGCCTCCTGCTGCCGTGGTCAGTTTGGGAGCCGGCATCGCGTCCGCCCAG
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  TCGGCCTCCTGCTGCCGTGGTCAGTTTGGGAGCCGGCATCGCGTCCGCCCAG

Qy   269  CTGCCCAGGAGGAGAGCTGGTGCAGAGAGGACCACCGGGACTAGTTCGGCCTCGCAGA
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  CTGCCCAGGAGGAGAGCTGGTGCAGAGAGGACCACCGGGACTAGTTCGGCCTCGCAGA

Qy   329  CAGAAGAAAGCCAGGATCCTGCGCCTTTCCTGAACGGCTCGAAGAGAGCTGAGATCC
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  CAGAAGAAAGCCAGGATCCTGCGCCTTTCCTGAACGGCTCGAAGAGAGCTGAGATCC

Qy   389  CACCTAAAGGCCGGAAAACACGGGCTCGAAGAGCTCGAGCCCATTATGAAGTT
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  CACCTAAAGGCCGGAAAACACGGGCTCGAAGAGCTCGAGCCCATTATGAAGTT

Qy   449  CACGACCTGGACAGGACGGAGCGCAGGCAGTGTGACGGACAGTGAGTGGCTGG
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  CACGACCTGGACAGGACGGAGCGCAGGCAGTGTGACGGACAGTGAGTGGCTGG

Qy   509  AAGCCAGAATCAACAGTCCAGCCTCCAGCCGTACAACCGGCTACAACCGGCCAGATCGGGAGTTT
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  AAGCCAGAATCAACAGTCCAGCCTCCAGCCGTACAACCGGCTACAACCGGCCAGATCGGGAGTTT

Qy   569  TCACCCGGGCTGGGCTCTACTACCTGTACTTGCACTTGCACTGTCTGGCCTGGCCTGCTGGAGG
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  TCACCCGGGCTGGGCTCTACTACCTGTACTTGCACTTGCACTGTCTGGCCTGCTGGAGG

Qy   629  TCTACCTGAAGTCTGAACTTGCTGGTGGATGTGTTGCTGGCCCTGCCTGGAGG
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  TCTACCTGAAGTCTGAACTTGCTGGTGGATGTGTTGCTGGCCCTGCCTGGAGG

Qy   689  TCTCAGCCACTGCGGCCCAGTTCCCTCGGGGCCCCAGCTCCGCCTTCGCCAGGTGTCT(
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  TCTCAGCCACTGCGGCCCAGTTCCCTCGGGGCCCCAGCTCCGCCTTCGCCAGGTGTCT(
```

```
                 ||||||||||||||||||||||||||||||||||||||
         CCGGGGCTGGGCTCTACTACCTGTACTGTCAGGTGCACTTTGATGATGAGGGGAAGGCTG  540

Qy    CCTGAAGCTGGACTTGCTGGTGGATGGTGTGCTGGCCCTGCGCTGCCTGGAGGAAT  688
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    CCTGAAGCTGGACTTGCTGGTGGATGGTGTGCTGGCCCTGCGCTGCCTGGAGGAAT  600

Qy    AGCCACTGCGGCCAGTTCCCTGCGGGCCCCAGCTCCGGCCTCTGCCAGGTGTCTGGGC  748
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    AGCCACTGCGGCCAGTTCCCTGCGGGCCCCAGCTCCGGCCTCTGCCAGGTGTCTGGGC  660

Qy    GGCCCTGCGCGCCAGGGTCCTCCCTGCGGATCCGCACCCTCCCCTGGGCCCATCTCA  808
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    GGGCCCTGCGGGCCAGGGTCCTCCCTGCGGATCCGCACCCTCCCCTGGGCCCATCTCA  720

Qy    TGCCCCTTCCTCACCTACTTCGGACTCTTCCAGGTTCACTGAGGGGCCCTGGTCT  868
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    TGCCCCTTCCTCACCTACTTCGGACTCTTCCAGGTTCACTGAGGGGCCCTGGTCT  780

Qy    ACAGTGCCCCAGGCTGCGGCCTCCCCTCGACAGCTCTCTGGGCACCGGTCCCT  928
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    ACAGTGCTCCCAGGCTGCGGCCTCCGGCTCTCTGACAGCTCTCTGGGCACCGGTCCCCT  840

Qy    CCCACCCTCAGCCGCTCTTTGCTCCAGACCTGCCCTCCCCTCTAGAGGCTGCCTGG  988
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    CCCACCCTCAGCCGCTCTTTGCTCCAGACCTGCCCTCCCCTCTAGAGGCTGCCTGG  900

Qy    GTTCACGTGTTTTCCATCCCACATAAATACAGTATTCCCACTCTTATCTTACACAT  1048
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    GTTCACGTGTTTTCCATCCCACATAAATACAGTATTCCCACTCTTATCTTACAACT  960

Qy    CACCGCCACTCTCCACCTCACTAGCTCCCAATCCCTGACACCCTTTGAGGCCCCCA  1108
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    CACCGCCCACTCTCCACCTCACTAGCTCCCAATCCCTGACCCTTTGAGGCCCCCA  1020

Qy    TCTCGACTCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTG  1168
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    TCTCGACTCCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTG  1080

Qy    AGGATGGTCCAGAAGACCCACTTCAGGCACTACAGAGGGGGTCTGGACCTGCGGCA  1228
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    AGGATGGTCCAGAAGACCCACTTCAGGCACTACAGAGGGGGTCTGGACCTGCGGCA  1140

Qy    GCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAG  1288
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    GCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAG  1200

Qy    GCTCCTCCCTTGAGAATTCCCTGTGATTTTTAAAACAGATATTTTTATTATT  1348
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    GCTCCTCCCTTGAGAATTCCCTGTGATTTTTAAAACAGATATTTTTATTATT  1260

Qy    TGACAAAATGTTGATAAATGG  1373
         |||||||||||||||||||||
Db    TGACAAAATGTTGATAAATGG  1285
```

lication US/10272411
JS2003010006A1
ON:
B Jewish Hospital
Jonathan
s, F. Patrick
elbaum, Steven
ON: RANKL MIMICS AND USES THEREOF
60019620-0202
ION NUMBER: US/10/272,411
DATE: 2002-10-15
ON NUMBER: 60/329,393
E: 2001-10-15
NOS: 52
In version 3.1

```
  89  CACAGCCCCCGCCCTCCCCATGGCGCGCCGTGGAGCCAGAGGCGGAGGGGCGCCC        808
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
   1  CAGCAGCCCCCGCCCTCCCCATGGCGCGCCGCCCGTGGAGCCAGAGGCGGAGGGGCGCCA  720

 149  AGCCGGGCACCGCCCTGCTGGTCCGCGCTGCGCCTGGGCCTGGGCCTGGCGCTGTGG        868
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  61  AGCCGGGCACCGCCCTGCTGGTCCGCGCTGCGCCTGGGCCTGGGCCTGGCGCTGTGG  780

 209  TCGGCCTTCCTGCTGGCCGTGGTCAGTTTGGGGAGCGCCGGCATCGCTGCGTCCGCCCT      928
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 121  TCGGCCTTCCTGCTGGCCGTGGTCAGTTTGGGGAGCGCCGGCATCGCTGCGTCCGCCCA  840

 269  CTGCCCAGGAGGAGCTGGTGGCAGAGGAGGAGGACCAGGACCCGTCGGAACTGAATCC      988
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 181  CTGCCCAGGAGGAGCTGGTGGCAGAGGAGGAGGACCAGGACCCGTCGGAACTGAATCC  900

 329  CAGAAGAAAGCCAGGATCCTGCGCCTTCCTGAACCGACTAGTTCGGCCTGCAC       1048
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 241  CAGAAGAAAGCCAGGATCCTGCGCCTTCCTGAACCGACTAGTTCGGCCTGCAG  960

 389  CACCTAAAGGCCGGAAAACACGGGCTCGAAGAGCGGATCGCAGCCCATTATGAAGT      1108
      |||||||||||||||||||||||||||||||||||||||||||||||||||||
 301  CACCTAAAGGCCGGAAAACACGGGCTCGAAGAGCGGATCGCAGCCCATTATGAAGT  1020

 449  CACGACCTGACACAGAGCGGAGCGCAGGCAGGTGTGGACGGGACAGTGAGTGGCTC      1168
      |||||||||||||||||||||||||||||||||||||||||||||||||||||
 361  CACGACCTGACACAGAGCGGAGCGCAGGCAGGTGTGGACGGGACAGTGAGTGGCTC  1080

 509  AAGGATGGTGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGCA      1228
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 421  AAGGATGGTGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGCA  1140

 569  TCACCCGGGCTGGGCTGGCTCTACTACCTGTACTTGTCAGTGCACTTTGATGAGGGGAA      1288
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 481  TCACCCGGGCTGGGCTGGCTCTACTACCTGTACTTGTCAGTGCACTTTGATGAGGGGAA  1200

 629  TCTACCTGAAGCTGGACTTGCTGGTGGATGGTGTGCTGGCCCTGCCTGGA       1348
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 541  TCTACCTGAAGCTGGACTTGCTGGTGGATGGTGTGCTGGCCCTGCCTGGA  1260

 689  TCTCAGCCACTGCGGCCAGTTCCTGCGGGCCCCAGCTCCGGCCTCGCCTCTGCCAGTGTC      1373
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 601  TCTCAGCCACTGCGGCCAGTTCCTGCGGGCCCCAGCTCCGGCCTCGCCTCTGCCAGTGTC  1285

 749  TGTTGGCCCTGCGCGGCCAGGGTCCTCCCTGCGGATCCGCACCCTCCCCTGGGCCCA
      ||||||||||||||||||||||||||||||||||||||||||||||||||||
 661  TGTTGGCCCTGCGCGGCCAGGGTCCTCCCTGCGGATCCGCACCCTCCCCTGGGCCCA

 809  AGGCTGCGCCCCTTCCTCACCTACTTCGGACTCTTCCAGTTCACTGAGGGGCCCT
      ||||||||||||||||||||||||||||||||||||||||||||||||||||
 721  AGGCTGCGCCCCTTCCTCACCTACTTCGGACTCTTCCAGTTCACTGAGGGGCCCT

 869  CCCCACAGTCGTCCCAGGCGTCCGCGCTCCCCTGACAGCTCTCTGGGCACCCGGT
      ||||||||||||||||||||||||||||||||||||||||||||||||||||
 781  CCCCACAGTCGTCCCAGGCGTCCGCGCTCCCCTGACAGCTCTCTGGGCACCCGGT

 929  CTGCCCACCCTCAGCCGCTCTTTGTCTCCAGCAGCCTGCCCCTCCCCTCTCTAGAGGCTG
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 841  CTGCCCCACCCTCAGCCGCTCTTTGTCTCCAGCAGCCTGCCCCTCCCCTCTCTAGAGGCTG

 989  GCCTGTTCACGTGTTTCCATCCCACATAAATACAGTATTCCCACTCTTATCTTA
      |||||||||||||||||||||||||||||||||||||||||||||||||
 901  GCCTGTTCACGTGTTTCCATCCCACATAAATACAGTATTCCCACTCTTATCTTA

1049  CCCCCACCGCCCACTCTCCACTCACTAGCTCCCCAATCCTGACCCTTTGAGGC
      |||||||||||||||||||||||||||||||||||||||||||||||||||
 961  CCCCCACCGCCCACTCTCCACTCACTAGCTCCCCAATCCTGACCCTTTGAGGC

1109  GTGATCTGACTCCCCCGGCCACGACCCCCAGGCATTGTGTTCACTGTACT
      ||||||||||||||||||||||||||||||||||||||||||||||||||
1021  GTGATCTGACTCCCCCGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACT

1169  GGCAAGGATGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCACTGGACCTGGC
```

Qy
Db
Qy
Db
Qy
Db
Qy
Db
Qy
Db
Qy
Db
Qy
Db
Qy
Db
Qy
Db
Qy
Db
Qy
Db
Qy
Db
Qy
Db
Qy
Db
Qy
Db
Qy

un

.arity   93.6%;   Score 1285;   DB 14;   Length 1306;
         100.0%;   Pred. No. 0;
onservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Db   421  AAGCCAGAATCAACAGCTCCAGCCTTCGCGCTACAACCGCCAGATCGGGGAGTT

Qy   569  TCACCCGGGCTGGCGTCGGCTCTACTACTGTACTGTCAGGTGCACTTTGATAGGGGAA
          ||||||||||||||||||||||||||||||||||||||||||||
Db   481  TCACCCGGGCTGGGCTCTCACTACTGTACTGTCAGGTGCACTTTGATGAGGGAA

Qy   629  TCTACCTGAAGCTGGACTTGCTGTGTGGATGGTGTGCTGGCCCTGCGCGCTGCCTGGA
          ||||||||||||||||||||||||||||||||||||||
Db   541  TCTACCTGAAGCTGGACTTGCTGTGTGGATGGTGTGCTGGCCCTGCGCGCTGCCTGGA

Qy   689  TCTTCAGCCACTGCGGCCAGTTCCCTCGGGCCCCAGCTCCGCCTCCGCCAGGTGTC
          |||||||||||||||||||||||||||||||||||||||||||
Db   601  TCTCAGCCACTGCGGCCAGTTCCCTCGGGCCCCAGCTCCGCCTCCGCCAGGTGTC

Qy   749  TGTTGGCCCTGCGGCGCCAGGTCCTCCCTGCGATCCGCACCCTCCCTGGGCCCA
          |||||||||||||||||||||||||||||||||||
Db   661  TGTTGGCCCTGCGGCGCCAGGGTCCTCCCTGCGATCCGCACCCTCCCTGGGCCCA

Qy   809  AGGCTGCCCCCTTCCTCACTACTTCGGACTCTTCCAGGTTCACTGAGGGGCCCTC
          ||||||||||||||||||||||||||||||||
Db   721  AGGCTGCCCCCTTCCTCACTACTTCGGACTCTTCCAGGTTCACTGAGGGGCCCTC

Qy   869  CCCCACAGTCGTCCCAGGCTGCCGGCTCCCGTCCCCGCACAGCTCTCTGGGCACCCGGT
          |||||||||||||||||||||||||||||||||||||||
Db   781  CCCCACAGTCGTCCCAGGCTGCCGGCTCCCGTCCCCGCACAGCTCTCTGGGCACCCGGT

Qy   929  CTGCCCCACCCTCAGCGCCGTCTTTGCTCCGACCTGCCCCTCCCTCTAGAGGCTGC
          |||||||||||||||||||||||||||||||||||
Db   841  CTGCCCCACCCTCAGCGCCGTCTTTGCTCCGACCTGCCCCTCCCTCTAGAGGCTGC

Qy   989  GGCTGTTCACGTGTTTTCCATCCACATAAATACAGTATTCCCATTCTTATCTTAC
          ||||||||||||||||||||||||||||||||||||
Db   901  GGCTGTTCACGTGTTTTCCATCCACATAAATACAGTATTCCCATTCTTATCTTAC

Qy  1049  CCCCACCGCCCACTCTCCACCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCC
          |||||||||||||||||||||||||||||||
Db   961  CCCCACCGCCCACTCTCCACCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCC

Qy  1109  GTGATCTGCACTCCCCCCTGGCCACAGACCCCACTTGTTCACTGTACTC
          ||||||||||||||||||||||||||||
Db  1021  GTGATCTGCACTCCCCCCTGGCCACAGACCCCACTTGTTCACTGTACTC

Qy  1169  GGCAAGGATGGGTCCAGAAGACCCACTTCAGGCACTAAGAGGGGCTGGACCTGGC
          ||||||||||||||||||||||||||||||||||||
Db  1081  GGCAAGGATGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGC

Qy  1229  GGAAGCCAAAGAGACTGGGGCTAGGGCAGGAGTTCCCAAATGTGAGGGGCGAGAAA
          ||||||||||||||||||||||||||||||||||
Db  1141  GGAAGCCAAAGAGACTGGGGCCTAGGGCAGGAGTTCCCAAATGTGAGGGGCGAGAAA

Qy  1289  ACAAGCTCCTCCCTGAGAATTCCCGTGGATTTTTAAAACAGATATTATTTTAT
          |||||||||||||||||||||||||||||||||||
Db  1201  ACAAGCTCCTCCCTGAGAATTCCCGTGGATTTTTAAAACAGATATTTTTTTAT

Qy  1349  ATTGTGACAAAATGTTGATAAATGG  1373
          |||||||||||||||||||||||||
Db  1261  ATTGTGACAAAATGTTGATAAATGG  1285
```

RESULT 8
US-10-310-793-29
; Sequence 29, Application US/10310793
; Publication No. US20030198640A1
; GENERAL INFORMATION:
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ni, Jian
; APPLICANT: Rosen, Craig A
; APPLICANT: Zhang, Jun
; APPLICANT: Wei, Ping
; TITLE OF INVENTION: Methods And Compositions For Treating Inflammat:
; TITLE OF INVENTION: Relating To Human Tumor Necrosis Factor-Gamma :
; FILE REFERENCE: PF573
; CURRENT APPLICATION NUMBER: US/10/310,793

```
          |||||||||||||||||||||||||||||||||||||||||||||
Db        AGGATGGTGGTCCAGAGAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCCGCA  1140

Qy        AGCCAAAGAGAGACTGGGGCTAGGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAG  1288
          |||||||||||||||||||||||||||||||||||||||||||||
Db        GCCCAAAGAGAGACTGGGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAG  1200

Qy        GCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTATTTTTATTATT  1348
          ||||||||||||||||||||||||||||||||||||||
Db        GCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTTTTTTATTATT  1260

Qy        TTGACAAAATGTTGATAAATGG  1373
          ||||||||||||||||||||||
Db        TTGACAAAATGTTGATAAATGG  1285
```

```
                    lication US/10272328A1
                    US2003010944A1
                    ION:
                    es Jewish Hospital
                    , Jonathan
                    s, F. Patrick
                    telbaum, Steven
                    ION: RANKL MIMICS AND USES THEREOF
                    60019620-0206
                    TION NUMBER: US/10/272,328A
                    DATE:   2003-01-24
                    ON NUMBER: 60/329,393
                    TE: 2001-10-15
                    D NOS: 51
                    tin version 3.1
```

sapiens

```
          93.6%;  Score 1285;  DB 14;  Length 1306;
arity     100.0%;  Pred. No. 0;
onservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

          |||||||||||||||||||||||||||||||||||
GCCCCCGGCCCCATGGCCGCCCGTCGGAGCGGAGGCGGAGGGGCGGCCCGGGGGGG  148
          ||||||||||||||||||||||||||||||||||
GCCCCGGCCCCCATGGCCGCCCGTCGGAGCCAGAGGCGGAGGGGCGGCCCGGGGGGG  60

GGGCACCGCCCCTGCTGTCCCCGCTCGGCGTCGGCCTGGCCGTGGCGGCCTGGCCTGCC  208
          |||||||||||||||||||||||||||||||||||
GGGCACCGCCCCTGCTGTCCCCGCTCGGCGTCGGCCTGGCCGTGGCCGCCTGGCCTGCC  120

CCTCCTGCTGGCCGTGGTCGGCAGTTTGGGGAGCGGGAGCGGGCATGCGGCTGTCCGCCCAGGAGC  268
          ||||||||||||||||||||||||||||||||||||
CCTCCTGCTGGCCGTGGTCAGTTTGGGGAGCGGGAGCGGGCATGCGGCTGTCCGCCCAGGAGC  180

CCAGGAGGAGGAGCTGGTGGCAGAGGAGGAGGAGGACCAGGAGGACCAGGAACTGAATCCCCAGA  328
          ||||||||||||||||||||||||||||||||
CCAGGAGGAGGAGCTGGTGGCAGAGGAGGAGGAGGACCAGGAGGACCAGGAACTGAATCCCCAGA  240

AGAAAGCCAGGATCCTGCGCCTTTCTGAACCGACTAGTTCGGCCTTCGCGCTCGCGAAGTG  388
          |||||||||||||||||||||||||||||||||||
AGAAAGCCAGGATCCTGCGCCTTTCCTGAACCGACTAGTTCGGCCTTCGCGCTTCGCGAAGTG  300

TAAAGGCCGGAAAACACGGGCTCGAAGAGGCGATCGCAGCCCATTATGAAGTTCATC  448
          ||||||||||||||||||||||||||||||||||
TAAAGGCCGGAAAACACGGGCTCGAAGAGCCGATCGCAGCCCATTATGAAGTTCATC  360

ACCTGACAGGAGCGGAGGCGCAGGCGCAGGTGTTGACGGGGACAGTGGCTGGCTGGGAGG  508
          |||||||||||||||||||||||||||||||||
ACCTGACAGGAGCGGAGGCGCAGGCGCAGGTGTTGACGGGGACAGTGGCTGGCTGGGAGG  420

CAGAATCAACAGCTCCAGCCTTCGCGCTACAACCGCCAGATCGGGGAGTTTATAG  568
          |||||||||||||||||||||||||||||
```

. DATE: 2002-12-06
ION NUMBER: 60/336,695
ATE: 2001-12-07
ION NUMBER: 10/226,294
ATE: 2002-08-23
ION NUMBER: 60/314,381
ATE: 2001-08-24
ION NUMBER: 09/899,059
ATE: 2001-07-06
ION NUMBER: 60/278,449
ATE: 2001-03-26
ION NUMBER: 60/216,879
ATE: 2000-07-07
ION NUMBER: 09/559,290
ATE: 2000-04-27
ION NUMBER: 60/180,908
ATE: 2000-02-08
ION NUMBER: 60/134,067
ATE: 1999-05-13
ION NUMBER: 60/132,227
ATE: 1999-05-03
r Application data removed - See File Wrapper or PALM..
ID NOS: 71
ntIn version 3.1

an

larity   93.6%;   Score 1285;   DB 14;   Length 1306;
Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
AGCCCCCGCGCCCCCATGGCCGCCCGTCGGAGCCCAGGAGGCGGAGGGGCGCCGGGGG 148
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGCCCCCGCGCCCCCATGGCCGCCCGTCGGAGCCCAGGAGGCGGAGGGGCGCCGGGGG 60

CGGGCACCGCCCCTGCTGGTCCCGCTCGGGCCTGGGCCTGCCGCCTGGCCTGCTGCC 208
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CGGGCACCGCCCCTGCTGGTCCCGCTCGGGCCTGGGCCTGCCGCCTGGCCTGCTGCC 120

GCCTCCTGCTGGCCGGTGGTCAGTTGGGAGGCCGGGCATCGTGTCCGCCCAGGAGGC 268
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCCTCCTGCTGGCCGGTGGTCAGTTGGGAGGCCGGGCATCGTGTCCGCCCAGGAGGC 180

CCCAGGAGGAGCCAGGACCTGCGGTGGGCAGGAGGACCAGGACCCGTCGGAACTGGAACTGGAACTGCCCAGA 328
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCAGGAGGAGCCAGGACCTGCGGTGGGCAGGAGGACCAGGACCCGTCGGAACTGGAACTGCCCAGA 240

AGAAAGCCAGGAGACTCGCGCCTTTCCTGAACCGACTAGTTCGGCCTCGCAGAAGTG 388
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGAAAGCCAGGAGACTCGCGCCTTTCCTGAACCGACTAGTTCGGCCTCGCAGAAGTG 300

TAAAGGCCGGAAAACACGGGCTCGAAGAGCGATCGCAGCCCATATGAAGTTCATC 448
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TAAAGGCCGGAAAACACGGGCTCGAAGAGCGATCGCAGCCCATATGAAGTTCATC 360

GACCTGGACAGGACCGGAGCGGAGCGCAGCGGTGTGGACGGACAGTGAGTGGCTGGGAGG 508
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GACCTGGACAGGACCGGAGCGGAGCGCAGCGGTGTGGACGGACAGTGAGTGGCTGGGAGG 420

CCAGAATCAACAGCTCCAGCCCTCTGCGCTACACCGCCAGATCGGGAGTTTATAG 568
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAGAATCAACAGCTCCAGCCCTCTGCGCTACACCGCCAGATCGGGAGTTTATAG 480

CCGGGCTGGGCTCTACTACCTGTACTGTCAGGTGCACTTGATGAGGGGAAGGCTG 628
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCGGGCTGGGCTCTACTACCTGTACTGTCAGGTGCACTTGATGAGGGGAAGGCTG 540

ACCTGAAGCTGGACTTGCTGGTGGATGGTGTGCTGGCCCTGCGCTGCGCTGGAGGAAT 688
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCTGAAGCTGGACTTGCTGGTGGATGGTGTGCTGGCCCTGCGCTGCGCTGGAGGAAT 600
```

```
Qy   689   TCTCAGCCACTGCGGCCAGTTCCTGGGCCCAGTTCCGCCTCTGCCAGGTGT(
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601   TCTCAGCCACTGCGGCCAGTTCCTGGGCCCAGTTCCGCCTCTGCCAGGTGT(

Qy   749   TGTTGGCCCTGGCCAGGGTCCTCCTCCCCTGCGCGGATCCCGCACCCTCCCTGGGCCCA
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661   TGTTGGCCCTGGCCAGGGTCCTCCTCCCCTGCGCGGATCCCGCACCCTCCCTGGGCCCA

Qy   809   AGGCTGCCCCTTCCTCACCTACTTCGGACTCTTCAGGTTCACTGAGGGGCCCT
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721   AGGCTGCCCCTTCCTCACCTACTTCGGACTCTTCAGGTTCACTGAGGGGCCCT

Qy   869   CCCCACAGTCGTCCCAGGCTGCGGCTCCCCTGGCCGACAGCTCTCGGGCACCCGGT
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781   CCCCACAGTCGTCCCAGGCTGCGGCTCCCCTGGCCGACAGCTCTCGGGCACCCGGT

Qy   929   CTGCCCCACCCTCAGCCGCTCTTGCTCCAGACCTGCCCCTCCCCTCTCCTAGAGGCTG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841   CTGCCCCACCCTCAGCCGCTCTTGCTCCAGACCTGCCCCTCCCCTCTCCTAGAGGCTG

Qy   989   GCCTGTTCACGTGTTTCCATCCACATAAATACAGTATTCCCACTCTTATCTTA
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901   GCCTGTTCACGTGTTTCCATCCACATAAATACAGTATTCCCACTCTTATCTTA

Qy   1049  CCCCACCGCCCACTTCTCCACCTCACTAGCTCCCAATCCCTGACCCTTTGAGGC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961   CCCCACCGCCCACTTCTCCACCTCACTAGCTCCCAATCCCTGACCCTTTGAGGC

Qy   1109  GTGATCTGACTCCCCCCTGGCCACAGACCCCAGGCATTGTGTTCACTGTACT
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  GTGATCTGACTCCCCCCTGGCCACAGACCCCAGGCATTGTGTTCACTGTACT

Qy   1169  GGCAAGGATGGTCCAGAGACACCCACTTCAGCGCACTAAGAGGGCTGGACCTGG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  GGCAAGGATGGTCCAGAGACACCCACTTCAGCGCACTAAGAGGGCTGGACCTGG

Qy   1229  GGAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGCGAGAA
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  GGAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGCGAGAA

Qy   1289  ACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTATTTTA
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201  ACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAACAGATATTATTTTA

Qy   1349  ATTGTGACAAAATGTTGATAAATGG 1373
           |||||||||||||||||||||||||
Db   1261  ATTGTGACAAAATGTTGATAAATGG 1285
```

RESULT 9
US-09-742-454A-1
; Sequence 1, Application US/09742454A
; Patent No. US2002004187A1
; GENERAL INFORMATION:
;   APPLICANT: WILEY, Steven R.
;   TITLE OF INVENTION: TWEAK Receptor
;   FILE REFERENCE: 2968-B
;   CURRENT APPLICATION NUMBER: US/09/742,454A
;   CURRENT FILING DATE: 2000-12-19
;   PRIOR APPLICATION NUMBER: 60/172,878
;   PRIOR FILING DATE: 1999-12-20
;   PRIOR APPLICATION NUMBER: 60/203,347
;   PRIOR FILING DATE: 2000-05-10
;   NUMBER OF SEQ ID NOS: 7
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 898
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (52)..(873)

; SEQ ID NO 1
; LENGTH: 898
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: TWEAK fusion protein construct
; NAME/KEY: CDS
; LOCATION: (52)..(873)
; OTHER INFORMATION:
US-09-883-777-1

Query Match 45.8%; Score 629.2; DB 9; Length 898;
Best Local Similarity 99.5%; Pred. No. 1.1e-162;
Matches 631; Conservative 0; Mismatches 3; Indels 0; 0

Qy 232 AGTTTGGGGAGCCGGGCATCGCTGTCCGCCCAGGAGGAGCTGGT
       |||||||||||||||||||||||||||||||||||||||||||||
Db 250 AGTTTGGGGAGCCGGGCATCGCTGTCCGCCCAGGAGGAGCTGGT

Qy 292 GAGGAGGACCAGGACCCGTCGGAACTGAATCCCAGACAGAAAGCCAGGATCC
       ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 310 GAGGAGGACCAGGACCCGTCGGAACTGAATCCCAGACAGAAAGCCAGGATCC

Qy 352 CCTTTCCTGAACCGACTAGTTCGGCTTCGGCTCGCAGAAGTGCACCTAAAGGCCGAAAAC
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 370 CCTTTCCTGAACCGACTAGTTCGGCTTCGGCTCGCAGAAGTGCACCTAAAGGCCGAAAAC

Qy 412 GCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGACAGGACGG
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 430 GCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGACAGGACGG

Qy 472 CAGGCAGGTGTGGACGGGACAGTGAGTGGCTGGGGAGGAGTTTATAGTCAACAAGCTC
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 490 CAGGCAGGTGTGGACGGGACAGTGAGTGGCTGGGGAGGAGTTTATAGTCAACAAGCTC

Qy 532 CCTCTGCCTACAACCGCCAGATCGGGAGTTTGATGAGGGGAAGGCTGTCTTACCTGGCTCTA
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 550 CCTCTGCCTACAACCGCCAGATCGGGAGTTTGATGAGGGGAAGGCTGTCTTACCTGGCTCTA

Qy 592 CTGTACTGTCAGGTGCACTTTGATGAGGGGAAGGCTGTCTTACCTGGCTGGACTT
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 610 CTGTACTGTCAGGTGCACTTTGATGAGGGGAAGGCTGTCTTACCTGGCTGGACTT

Qy 652 GTGGATGGTGTGCTGGCCCTGCGCTGCCTGGAGAATTCTCAGCCACTGCCGGCCAG
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 670 GTGGATGGTGTGCTGGCCCTGCGCTGCCTGGAGAATTCTCAGCCACTGCCGGCCAG

Qy 712 CTCCGGGCCCCAGCTCCGCCTCCTGGCCACCCTCCCTGTTGGCCCTGTCGGGCCTGG
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 730 CTCGGGGCCCCAGCTCCGCCTCCTGGCCACCCTCCCTGTTGGCCCTGTCGGGCCTGG

Qy 772 TCCCTGCGGATCCGACACCTCCCCTGGGCCCATCTCAAGGCTGCCCCTTCCTCAC
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 790 TCCCTGCGGATCCGACACCTCCCCTGGGCCCATCTCAAGGCTGCCCCTTCCTCAC

Qy 832 TTCGGACTCTTCCAGGTTCACTGAGGGGCCCTGG 865
       ||||||||||||||||||||||||||||||||||
Db 850 TTCGGACTCTTCCAGGTTCACTGAGGGGCCCTGG 883

ION: Description of Artificial Sequence: human TWEAK
ION: fusion protein construct

arity 45.8%; Score 629.2; DB 9; Length 898;
99.5%; Pred. No. 1.1e-162; Indels 0; Gaps 0;
onservative 0; Mismatches 3;

Qy TGGGGAGCCGGGCATCGCTGTCCGCCCAGGAGGAGCTGCCCAGGAGGAGCTGGTGGCA 291
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db TGGGGAGCCGGGCATCGCTGTCCGCCCAGGAGGAGCTGCCCAGGAGGAGCTGGTGGCA 309

Qy AGGACCAGGACCCGTCGGAACTGAATCCCAGACAGAAGAAGCCAGGATCCTGCG 351
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db AGGACCAGGACCCGTCGGAACTGAATCCCAGACAGAAGAAGCCAGGATCCTGCG 369

Qy TCCTGAACCGACTAGTTCGGCTCGCAGAAGTGCACCTAAAGGCCGGAAAAACACGG 411
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db TCCTGAACCGACTAGTTCGGCTCGCAGAAGTGCACCTAAAGGCCGGAAAAACACGG 429

Qy GAAGAGGCGATCGCAGCCCATTATGAAGTTCATCCACGACAGGACGAGCG 471
   ||||||||||||||||||||||||||||||||||||||||||||||||||
Db GAAGAGGCGATCGCAGCCCATTATGAAGTTCATCCACGACAGGACGAGCG 489

Qy CAGGTGTGGACGGGACAGTGAGTGGCTGGGGAGGAAGCCAGGAATCAACAGCTCCAGC 531
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db CAGGTGTGGACGGGACAGTGAGTGGCTGGGGAGGAAGCCAGGAATCAACAGCTCCAGC 549

Qy TGCGCTCAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTCTACTAC 591
   ||||||||||||||||||||||||||||||||||||||||||||||||||
Db TGCGCTCAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTCTACTAC 609

Qy ACTGTCAGGTGCACTTTGATGAGGGGAAGGCTGTCTACTGAAGCTGGACTTGCTG 651
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ACTGTCAGGTGCACTTTGATGAGGGGAAGGCTGTCTACTGAAGCTGGACTTGCTG 669

Qy ATGGTGTGCTGGCCCTGCGCTGCCTGGAGAATTCTCAGCCACTGCGGGCCAGTTCC 711
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ATGGTGTGCTGGCCCTGCGCTGCCTGGAGAATTCTCAGCCACTGCGGGCCAGTTCC 729

Qy GGCCCCAGCTCCGCCTCGCCAGGTGTCTGGGCTGTTGGGCCTGCCGGCCTGCCCAGGGTCC 771
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db GGCCCCAGCTCCGCCTCGCCAGGTGTCTGGGCTGTTGGGCCTGCCGGCCTGCCCAGGGTCC 789

Qy CGCGGATCCGACACCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACCTAC 831
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db CGCGGATCCGACACCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACCTAC 849

Qy GACTCTTCCAGGTTCACTGAGGGGCCCTGG 865
   ||||||||||||||||||||||||||||||
Db GACTCTTCCAGGTTCACTGAGGGGCCGCGG 883

cation US/09883777
0110853A1
ON:
, Steven R.
ON: TWEAK RECEPTOR
2968-C
ION NUMBER: US/09/883,777
DATE: 2001-06-18
DN NUMBER: US 60/172,878
E: 1999-12-20
N NUMBER: US 60/203,347
E: 2000-05-10
DN NUMBER: PCT/US00/34755
E: 2000-12-19
N NUMBER: US 09/742,454
E: 2000-12-19
NOS: 16
In version 3.1

ID NOS: 38054
SEQ for Windows Version 3.0

io sapiens

c_feature
..(493)
TION: n = A,T,C or G
25

larity   32.4%;   Score 445.4;   DB 10;   Length 493;
Conservative   99.8%;   Pred. No. 3.1e-112;   Indels   1;   Gaps   0;

Matches 346;   Conservative   0;   Mismatches   45;   Indels   8;

Qy   83 GGCAGGCACAGCCCCGCCCCCATGGCCGCCCGTCCGGAGCCAGAGGCGGAGGGC
              |||||||||||||||||| |||||||||||||||||| ||||||||||||||||||
Db   4 GGCAGGCACAGCCCC--GCCCCATGGCCACCCGTCGGATCCAGAGGCGGAGGGA

Qy   143 GGGGGGAGCCGGGCACCGCCCTCGTCCCGTCCGCGCCTGGCCTGGGCCTCGGC
              ||||||||||||||||||||||| |||||||||||||||| |||||||||||||||
Db   62 GGTGGGAGCCGGGCACCGCCCCGTTGTCGCTGCCTGGGCCTGGGCCTGGT

Qy   203 CCTGCCTCGGCCTCTGCTGCCGCGTGTCAGTTTGGGAGCCGGGCATCGCTGTC
              |||||||||||||||||||||||||||||| ||||||||||||||||||
Db   122 CCTGCCTCGGCCTCTGCTGCCGCGTGTCAGCTGGGGAGCCGGGCATCGCTGTT

Qy   262 --CAGGAGCCTGCCCAGGAGGAGCTGGTGGTGCAGGAGGAGGACCAGGACCCGTCGGA
              |||||||||||||||||||| ||||||||||||||||||||||||||
Db   182 AGCAGGAGCCTTCCCAGAGGGACCTGGTGGCAGAGAAGACCCGGACCCGCTGGA

Qy   320 ATCCCCAGACAGAAGAAAGCCAGGATCCTGCGCCTTTCCTGAACCGACTAGTTCG
              ||||||||||||||||||||||||| |||||||||||||||| || ||||||||
Db   242 ATCCCCAGCCGGAGGAAAGCCAGGATGCTGTTCCTTTCCTGAA---ACTGGTTCG

Qy   380 GCAGAAGTGCACCTAAAGGCCGGAAAACACGGGCTCGAAGAGCGATCGCAGCCCA
              |||||||||||||||||||||||||||||||||||||||||||||||
Db   299 GCAGAAATGCATCTAAAGGCCAGAAACCACCGGCGCGCAGAGCAATGCAGCCCA

Qy   440 AAGTTCATCCACGACCTGGACACAGGACGGAGGCGCAGGCAG 478
              ||||||||||||||||||||||||||||||||||||||
Db   359 AAGTTCGCCCACGACCACCAGGACAAGATGGAGCACAGGCAG 397

RESULT 13
US-10-085-783A-55176
; Sequence 55176, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;   APPLICANT: ChondroGene Inc.
;   APPLICANT: Liew, C.C.
;   TITLE OF INVENTION: Compositions and Methods Relating to Osteoart
;   FILE REFERENCE: 4231/2002
;   CURRENT APPLICATION NUMBER: US/10/085,783A
;   CURRENT FILING DATE: 2002-02-28
;   PRIOR APPLICATION NUMBER: US 60/305,340
;   PRIOR FILING DATE: 2001-07-13
;   PRIOR APPLICATION NUMBER: US 60/275,017
;   PRIOR FILING DATE: 2001-03-12
;   PRIOR APPLICATION NUMBER: US 60/271,955
;   PRIOR FILING DATE: 2001-02-28
;   NUMBER OF SEQ ID NOS: 58994
;   SOFTWARE: PatentIn version 3.2
; SEQ ID NO 55176
;   LENGTH: 213
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (212)..(213)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-55176

Query Match   13.7%;   Score 188.4;   DB 12;   Length 213;
Best Local Similarity 99.5%;   Pred. No. 1.1e-41;
Matches 189;   Conservative   0;   Mismatches   1;   Indels   0;   C

Qy 1184 AGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGCGGCAGGAAGCCAAAC
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 AGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGCGGCAGGAAGCCAAAC

Qy 1244 TGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAAACAAGAGCTCCTC
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TGGGCCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAAACAAGAGCTCCTC

Qy 1304 GAGAATTCCCTGTGGATTTTAAAACAGATATTATTTTATTATTGTGACAAA

ACTTCTTCCAGGTTCACTGAGGGGCCCTGGTCTCCCCACAGTCGTCCCAGGCTGCCGG 893
||||||||||||||||||||||||||||||||||||||||||||||||
ACTTCTTCCAGGTTCACTGAGGGGCCCCTGGTCTCCCCGCAGTCGTCCCAGGCTGCCGG 106

CCCTCGACAGCTTCTCTGGGCACCCGGTCCCCTGCCCCTCAGCGCGCTCTTTG 953
|||||||||||||||||||||||||||||||||||||||||||||||||||
CCCTCGACAGCTTCTCTGGGCACCCGGTCCCCTGCCCCTCAGCGCGCTCTTTG 166

CAGACCTGCCCCTCCCTCTAGAGGCTGCCTGGGCCTGTTCACGTGTTTTCATCCCA 1013
|||||||||||||||||||||||||||||||||||||||||||||||||
CAGACCTGCCCCTCCCTCTAGAGGCCTGCCTGGGCCTGTTCACGTGTTTTCATCCCA 226

AAATACAGTATTCCCACTCTTATCTTACAACTCCCCCACCGCCCACTCTCCACCTCA 1073
|||||||||||||||||||||||||||||||||||||||||||||||
AAATACAGTATTCCCACTCTTATCTTACAACTCCCCCACCGCCCACTCTCCACCTCA 286

GCTCCCCAAATCCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCCCCCTGGCCAC 1133
||||||||||||||||||||||||||||||||||||||||||||||||
GCTCCCCAATCCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCCCCCCTGGCCAC 346

CCCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCCC 1193
|||||||||||||||||||||||||||||||||||||||||||||||
CCCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCCC 406

TCAGGCCACTAAGAGGGGCTGGACCTGGCGCGGCAGGAAGCCAAAGAGACTCGGCCTAGG 1253
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCAGGCCACTAAGAGGGGCTGGACCTGGCGCGGCAGGAAGCCAAAGAGACTCGGCCTAGG 466

GGAGTTCCCAAATGTGAGGGGCGA 1280
||||||||||||||||||
GGAGTTCCCAAATGTGAGGGGCGA 493

Application US/09960352
2013 7139A1
ION:
en, Wesley C.
, Nengbing
tt, John C.
hialagan, Nagappan
TION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
TION: MUSCLE AND FAT DEPOSITION
16511.006/37-21(10298)C
TION NUMBER: US/09/960,352
DATE: 2001-09-24
D NOS: 15112

taurus
ION: Clone ID: 10-LIB3057-025-Q1-K1-C5

arity   20.5%;   Score 281;   DB 9;   Length 408;
86.7%;   Pred. No. 4.9e-67;

```
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ATTCCCTGTGGATTTTAAAACAGATATTATTTTATTATTATTGTGACAAAATGT 180
```

AAATGG 1373
|||||
AAATGG 190

76
Application US/10242535A
US20040013663A1
ION:
3roGene Inc.
W, C.C.
ION: Compositions and Methods Relating to Osteoarthritis
4231/2005
TION NUMBER: US/10/242,535A
DATE: 2002-09-12
DN NUMBER: US 10/085,783
TE: 2002-02-28
DN NUMBER: US 60/305,340
TE: 2001-07-13
DN NUMBER: US 60/275,017
TE: 2001-03-12
DN NUMBER: US 60/271,955
TE: 2001-02-28
O NOS: 58994
:In version 3.2

feature
..(213)
ON: n is a, c, g, or t
76

urity 13.7%; Score 188.4; DB 15; Length 213;
99.5%; Pred. No. 1.1e-41;
onservative   0; Mismatches   1; Indels   0; Gaps   0;

```
3ACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGCCAGGAAGCCAAAGAGAC 1243
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3ACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGCCAGGAAGCCAAAGAGAC 60

CTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTT 1303
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTT 120

TTCCCTGTGGATTTTAAAACAGATATTATTTTTATTATTATTGTGACAAAATGT 1363
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
TTCCCTGTGGATTTTAAAACAGATATTATTTTTATTATTATTGTGACAAAATGT 180
```

AAATGG 1373
|||||
AAATGG 190

plication US/09983965
0137160A1
ON:
n, Wesley C.
  Nengbing
t, John C.
ialagan, Nagappan
ON: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
ON: MUSCLE AND FAT DEPOSITION
37-21(10297)C
ION NUMBER: US/09/983,965

; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 2183
; LENGTH: 264
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; OTHER INFORMATION: Clone ID: 42-LIB3057-025-Q1-K1-C6
US-09-983-965-2183

Query Match   6.7%; Score 91.4; DB 9; Length 264;
Best Local Similarity   66.5%; Pred. No. 5.9e-15;
Matches 147; Conservative   0; Mismatches 71; Indels   3;   G

```
Qy  123 CCAGAGGCGGAGGGGGGCGCCGGGGGAGCCGGGCACCGGCCCTGCTCGTCCCCGCTCG
            ||||||| |||||||| ||||| |||||||| ||||| | |||||||||||||| |
Db   40 CCAGAGACGAGCGAGCGGGGGCGGTTCCGAGCCTGCCACGGGTCTGCTGACCAAGGTCT

Qy  183 GGGCCTGGGCTGGGCGGCTGGCCTGCCTCGGCCTCCTGCTGCTGCCGTGGTCAGTTTGG
            || |||||| ||| ||||| ||||| ||||| |||||||| | ||||| ||| ||
Db  100 GGACCCTGACCCTGGTGCTGACGTGCGTCGGCGTCGTCGTGCGGCCTGGCCTGCCAGGCTGG

Qy  243 CCGGGCATCGTCGTCCGCCCA---GGAGCCTGCCCAGGAGGAGGAGCTGGTGGCAGAGG
            ||||| | |||||||  |||| |||| |||||||
Db  160 CCGCGAATGCTGTTTGCCCAGACGACGGAGCCTGACCAGATATACCTGGTCAAACAAG

Qy  300 CCAGGACCCGTCGGAACTGAATCCCCAGACAGAAGAAAGCC 340
            || ||||| || |||  ||||| ||||||  |||
Db  220 CCGCGACGCGTGTCGATCTGAATCCCTGGGCGGGGAAATGAC 260
```

Search completed: April 8, 2004, 06:15:20
Job time : 530.048 secs

us-09-245-198a-4.rpr

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

.n search, using sw model

cil 7, 2004, 17:41:57 ; Search time 15.6228 Seconds
(without alignments)
1748.625 Million cell updates/sec

-09-245-198A-4
14
MSLLDFEISARRLPLPRSLG.........PWAHLKAAPLTYFGLFQVH 284

DSUM62
Dop 10.0 , Gapext 0.5

1366 seqs, 96191526 residues

s satisfying chosen parameters: 283366

th: 0
th: 2000000000

nimum Match 0%
ximum Match 100%
sting first 45 summaries

R 78:*
pir1:*
pir2:*
pir3:*
pir4:*

the number of results predicted by chance to have a
than or equal to the score of the result being printed,
d by analysis of the total score distribution.

SUMMARIES

ry ch	Length	DB	ID	Description
.4	254	2	I38427	4-1BB ligand - hum
.2	1323	2	S27224	N-methyl-D-asparta
.0	375	2	A75264	hypothetical prote
.9	378	2	H87333	HlyD family secret
.9	878	2	T17245	hypothetical prote
.8	206	2	T34961	probable membrane
.7	441	2	S41710	mitosis-specific c
.7	776	2	T36946	probable cation-tr
.5	1657	2	T15838	hypothetical prote
.4	310	2	D70745	hypothetical prote
.4	351	2	B34768	ORF5 protein - Orf
.4	566	2	T35203	probable two-compo
.3	212	2	A70611	hypothetical prote
.3	664	2	D83231	hypothetical prote
.3	681	2	H83044	2,4-dienoyl-CoA re
.3	1217	2	T00270	hypothetical prote
.3	814	2	G02878	cadherin-15 precur
.3	1329	2	D87226	conserved hypothet
.2	660	2	T03038	probable inhibitor
.2	886	2	S07132	hypothetical prote
.1	210	2	D87394	hypothetical prote
.1	278	2	A49266	fas ligand - rat
.1	281	2	I38707	Fas ligand - human
.1	998	2	G83022	probable two-compo
.1	531	2	C83153	conserved hypothet
.1	492	2	A87471	hypothetical prote
.0	197	2	JH0309	tumor necrosis fac
.0	204	1	S24641	lymphotoxin - bovi
.0	445	2	T30604	hypothetical prote
30	87	6.0	H83357	2 2124 probable nc
31	86.5	6.0	F83633	2 399 hypothetica
32	86.5	6.0	B70629	2 439 hypothetica
33	86.5	6.0	B75346	2 755 probable co
34	86.5	6.0	E98121	2 762 hypothetica
35	86.5	6.0	S14113	1 764 1-phosphati
36	86.5	6.0	D95256	2 810 ATP-depende
37	86.5	6.0	AG0535	2 887 ClpB-like p
38	86	6.0	B69953	2 139 hypothetica
39	86	6.0	G98331	2 565 probable ol
40	86	6.0	E83325	2 568 probable ch
41	86	6.0	A87309	2 777 hypothetica
42	86	6.0	P75518	2 839 hypothetica
43	86	6.0	B48349	1 933 glycoprotei
44	85.5	5.9	AE3059	2 563 hypothetica
45	85.5	5.9	C98227	2 563 hypothetica

ALIGNMENTS

RESULT 1
I38427
4-1BB ligand - human
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 21-Jul
C;Accession: I38427
R;Alderson, M.R.; Smith, C.A.; Tough, T.W.; Davis-Smith, T.; Armitage,
Eur. J. Immunol. 24, 2219-2227, 1994
A;Title: Molecular and biological characterization of human 4-1BB and i
A;Reference number: I38426; MUID:94374434; PMID:8088337
A;Accession: I38427
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-254 <RES>
A;Cross-references: EMBL:U03398; NID:g571322; PIDN:AAA53134.1; PID:g571

Query Match   7.4%;   Score 106.5; DB 2; Length 254;
Best Local Similarity 26.5%; Pred. No. 0.32;
Matches   74; Conservative 31; Mismatches 93; Indels 81; Ga

Qy   32 PPAPMAARRSQRRRGRRGEPGTALLVPLAIGLGLALACLGLLLAVVSL-GSRASL-S
          |||| : ||||
Db   16 PPAP-------RARACRVLP-WALVAGLLLLLLLAAACAVFLACPWAVSGARASPGS

Qy   90 PAQEELVAEEDQDPSELNPQTESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHY
          |||:| :
Db   68 PRLRE-----GPELSP-----DDPAGLLDL---------------RQGMFAQLVAQNV

Qy   150 PRPGQDGAQAGVDGTVSGWEB---ARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVH
                      :|: :| :|: : :|| ||: :   : : |||:||
Db   103 ------------IDGPLSWYSDPGLAGVSLTGGLSYKEDTKELVVAKAGVYYVVFFQ--

Qy   207 GKAVYLKLDLLVDG-------VLALRCLEEFSATAASSLGPQLRLCQVSG--------
          |: ||| :
Db   147 -----LELRRVVAGEGSGSVSLALHLQPLRSAAGAAALAITVDLPPASSEARNSAFG

Qy   250 -LLALRPGSSLRIRTLPWAHLKAAPFL----TYFGLFQV 283
           ||| | :| |||:|
Db   202 RLLHLSAGQRLGVHLHTEARARHAWQLTQGATVLGLFRV 240

RESULT 2
S27224
N-methyl-D-aspartate receptor epsilon-4 chain - mouse
C;Species: Mus musculus (house mouse)
C;Date: 25-Feb-1994 #sequence_revision 01-Sep-1995 #text_change 17-Mar
C;Accession: S27224
R;Ikeda, K.; Nagasawa, M.; Mori, H.; Araki, K.; Sakimura, K.; Watanabe,
FEBS Lett. 313, 34-38, 1992
A;Title: Cloning and expression of the epsilon-4 subunit of the NMDA re
A;Reference number: S27224; MUID:93050214; PMID:1385220
A;Accession: S27224

ary
mRNA
B; <IKE>
B: EMBL:D12822
nethyl-D-aspartate receptor 2D; glutamate receptor homology
nembrane protein
glutamte receptor homology <GRH>

larity   7.2%;   Score 104;   DB 2;   Length 1323;
          23.7%;   Pred. No. 3.2;
onservative   21;   Mismatches   97;   Indels   62;   Gaps   8;

LPLPRSLGSRDGGAVRQAQPPAPM-----AARRSQRRRGRRGEPGTALLV------ 57
  |||| |||||||| || |: | :|||| | |:|||
PPPQQPLPSPAYPAARPPPGPAPFVPRERAAADRWRRAKG-TGPPGGAALADGFHRY 961

--PLALGLGLALAC-LGLLLAVVSLGSRASLSAQEPAQEELVAEEDQDPSE------ 105
    |||| |:||| |: | | | :| |||||| |:
EPQGLGLGEARAAPRG- AAGRPLSPPTTQPPQKPPPSYFAIVREQEPAEPPAGAF 1019

------------LNPQT----EESQDPAPFLNRLVRPRRS-------------A 130
            |: |||||||||||| ||:   :|
SPPAPAAAAAVGPPLCRLAFEDESPPAPSAGRVLTPRASRCWVGARAARARALGPR 1079

KTRARRAIAAHYEVHPRPEGQDGAQAGVDGTVSGWEEARINSSSPLRYNRQIG 186
 :| | : : :: || ||| | |: :| || | |: |:|
RVRTAPPPCAYLDLEPSPSDSEDSESLGGASLGGLEPWWFADFPYPYAERLG 1135

ain - Deinococcus radiodurans (strain R1)
occus radiodurans
9 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
4

, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
mathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
enter, J.C.; Fraser, C.M.
1577, 1999
quence of the radioresistant bacterium Deinococcus radiodurans R1.
: A75250; MUID:20036896; PMID:10567266
4
ary
NA
<WHI>
: GB:AE002081; GB:AE000513; NID:g6460337; PIDN:AAF12062.1; PID:g646034
rce: strain R1

arity   7.0%;   Score 100.5;   DB 2;   Length 375;
          24.0%;   Pred. No. 1.5;
onservative   27;   Mismatches   84;   Indels   107;   Gaps   13;

CAQPPAPMAA-----RRSQRRGRRGEPGT--ALLVPLALGLGLALACLGLLLAVV 77
 | || || || || || | | ||:|||| |||||| | ||||
RCQPTPGRAATARRLRRPGRAPGRRGQPGSHPAARQRLRADLGAALA------LVA 58

RASLSAQE-PAQEELVAEEDQDPSELNPQTESQDPAPFLNRLVRPRRSAPKGRKT 136
 |||| | ||||||||| : ||
RPTLPAGELPTLLDLAR------------QDLESLDDSP-------------- 89

AIAAHYEVHPR-PGQDGAQAG--VDGTVSGWEEARINSSSPLRYNRQIGE-----F 188
|||||||| | ||| |||| |||||||||||: |:|||:|:||
AVQARAATFPRLPGDPGAGFGHPISGTLAGLE--RLSAASLRAHWARFGQRGSVLG 147

AGLYLYLY---------------CQVHFDEGKAVYLKLDLLVDGVLALRCLEEFS 231
   :| :| ||:|||||| | | |: || |:| | | |
ADAQEVYELVAGLFADWQPGEDRPMPAHFQPGLRLHL------------------ 188

SSLGPQLRLCQVS---------GLLALRPGSSLRIRT 263
|: | | || |: |: |

PGTALLVPLALGLGLALACLGLLLLAVVSLGSRASLSAQEPAQEELVAEEDQDPSEL 106
  : |:||||| :
TGDTVLLPLAQGGHRPLS------RAQSPAAPASLSALSSETPART 441

-----------EESQDP--------APFLNRLVRPRRSAPKGRKTRAR 139
  : |:| :
TGLIYDSVMLKHQCSCGDNSRHPEHAGRIQSIWSRLQERGLRSQCECLRGRKASLE 501

AHYEVH 149
VHSERH 511

protein - Streptomyces coelicolor
myces coelicolor
9 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
1
Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
MBL Data Library, August 1999
: Z21563
1
ary; translated from GB/EMBL/DDBJ
NA
<SAU>
: EMBL:AL109663; PIDN:CAB52000.1; GSPDB:GN00070; SCOEDB:SC4A10.24C
rce: strain A3(2)

A10.24c

rity   6.8%; Score 98; DB 2; Length 206;
27.3%; Pred. No. 1.2;
nservative 18; Mismatches 64; Indels 46; Gaps 7;

GAVRQAQPPAPMAARRSQRRRGR-----------------RGEPGTA---LLVPLALGL 63
  : :|
GKI-QTTPRGGRVSRKPELKGRAARLARLLPTGSGSRGQAARAPFVLLVVLLGG 64

ACLGLLLLAVVSLGSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQDPAPFLNRL 123
  : |||:
--IGLLVL----------------NSALSEGSFQLDDLKQRTKELTDEEQALQRD 103

RSAPKGRKTRARATAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPL 179
  :: ||
SAPRALQRRAR-------ELGMVPGGDPAFLDPDGTVKGVPSPAPAAATPL 151

clin 2 - garden snapdragon
num majus (garden snapdragon)
#sequence_revision 13-Jan-1995 #text_change 16-Jul-1999

n, E.S.; Murphy, G.J.P.; Doonan, J.H.
, 1994
f cell division revealed by transcriptional regulation of genes durin
S41709; MUID:9148008; PMID:8313906

cid sequence not shown
NA
FOB>
EMBL:X76123; NID:g425262; PIDN:CAA53729.1; PID:g425263

in
cle control; cell division control; mitosis

rity   6.7%; Score 97; DB 2; Length 441;
24.0%; Pred. No. 3.4;
nservative 44; Mismatches 104; Indels 74; Gaps 15;

Qy 25 GAVRQAQPPAPMAARRSQRR-----------------RGRRGEPGTALLVPLALGLGLAL
Db 19 GAIKQKS----MAVEKNRRALGDIGNVVTVRGVEGKALPQVSRPITRGF-----(

Qy 74 -------------LAVVSLGSRASLS----AQEPAQEELVAEEDQDPSELNPQ(
Db 70 ANAEAAAAENNKNSLAVNAKGADGALPIKRAVARVPVQKKTVKSKPQEIIEISPD

Qy 115 DPAPFLNRLVRPRRS---APKGRKTRARATAAHYEVHPRPGQ------DGAQAG\
Db 130 --APVLEKEITGEKSLKKKAPTLTSTLTARSKAASV-VRTKPEQIVIDAADVN\

Qy 166 SGWEE--------ARINSSSPLRY---NRQIGEFIVTRAGLYYLYCQVHFD---EG\
Db 187 VEYVEDMYKFYKSAENDSRPHDYMDSQPEINEKM--RAILIDWLVQVHYKFELSPF

Qy 213 KLDLLVDGVLALRC-----LEEFSATAASSLGPQLRLCQVSGLLALRPGS 257
Db 245 TIN-IVDRYLASKITSRRELQLLGMSSMLIASKYEETWAPEVNDLVCISDGS 295

RESULT 8
T36946
probable cation-transporting ATPase - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug
C;Accession: T36946
R;Seeger, K.J.; Harris, D.; Thomson, N.R.; Parkhill, J.; Barrell, B.G.
submitted to the EMBL Data Library, September 1999
A;Reference number: Z21607
A;Accession: T36946
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-776 <SEE>
A;Cross-references: EMBL:AL109962; PIDN:CAB53131.1; GSPDB:GN00070; SCO
A;Experimental source: strain A3 (2)
C;Genetics:
A;Gene: SCOEDB:SCJ1.13
C;Superfamily: ATPase nucleotide-binding domain homology
F;442-585/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match   6.5%; Score 93.5; DB 2; Length 776;
Best Local Similarity 24.1%; Pred. No. 12;
Matches   63; Conservative   30; Mismatches 103; Indels 65; G

Qy 47 RRGEPGTALLVPLALGLGLALA--CLGLLLA-VVSLGSRASLSAQEPAQEELVAEE
Db 73 RRGHAGVDLIAVLALGGTLAVGEYLAGVLIALMLATGRTLEGAAQRRASHDLHALL

Qy 104 SELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQDGAQA\
Db 133 RSARRRTGDG--------VVR----VPLSEITAGDALVVGPGEVVP---------

Qy 164 TVSGWE--EARINSSPLRYNRQIGE----FIVTRAGLYYLYCQVHFDEGKAVVL:
Db 170 RVESTEAVLDESVLTGEPLQVTRQREGEARSGAVNAGGAFDL--------RATAI

Qy 217 LVDGVLALRCLEEFSATAASSLGPQLRLCQ-----VSGLLALRPGSSLR-
Db 221 TYAGIVRL-----AQQAGAESAPVVRLADRYAAWFLPLALATAALAWLVSGSAVRJ

Qy 261 -IRTLPWAHLKAAPFLTYFGL 280
Db 275 LVVATPCPLLAARPVAVVSGL 295

RESULT 9
T15838
hypothetical protein C54D2.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Mar-
C;Accession: T15838
R;Minx, P.

EMBL Data Library, October 1995
e sequence of C. elegans cosmid C54D2.
r: Z18415
38

nary; translated from GB/EMBL/DDBJ
DNA
7 <MIN>
s: EMBL:U37548; NID:g1017804; PID:g1017809; PID:AAA79201.1; CESP:C54D2

2.5
67/3; 86/3; 121/2; 199/3; 230/2; 308/2; 334/2; 370/2; 439/3; 470/3; 513
dium channel protein

larity   6.5%;   Score 93.5;   DB 2;   Length 1657;
Conservative 23.2%; Pred. No. 29;   Indels 71;   Gaps   12;
             38;   Mismatches 99;

LLVPLALGLGLALACLG-----LLLAVVSLGSRAS------LSAQEPAQEELVAEED 100
        :|: |:|:: :   :|| ||||
QTNPWAALYFVALMTFGNYVLFNLLVAILVEGFQESKEEEKRQLEEDARKQAVEEED 940

SELNPQTEESQDPAPFLNRLVRP------RRSAPKGRKTRARRAIAAHYEVHPRPGQ 154
     :||:   :::|:::: ::|:| |::|:|.|||
RELELIIAKTTSPA--FNNGVAPAECTCQRPSSPE-ESPSPRLLSANY--HPSPER 994

QAGVDGTVSGWEEARINSSSPL------------RYNRQIGEFIVTRAGLYYLYCQV 202
     :|:   :|:| |:| ::| :||:||||||
-ANLDAIID--KRLVLRNSAPFDRSPVSEGRDDSRLNRHASLVLPVANGVPYRRQRV 1051

------------FDEGKAVYLKLDLLVDGVLALRCLE--EFSATAASSLGPQLRLCQV 247
            :| |::: :   :||| ::|: ||||||:||||
KASQELKQALAEEEKRNEAKQNTFVRKLLKKTCLHNRTEFS--------------- 1095

LALRPGGSLRIRTLPWAHLKAAPFLTYF 278
    :| :
FLMGPKNPLRIKCLQTTQKKWFDYTVLF 1124

ain Rv0497 - Mycobacterium tuberculosis (strain H37RV)
terium tuberculosis
98 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
05
sch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
tes, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rogers, J.; Rutter, S.; Seeger, K.; Skelton, J.; Squares, S.
4, 1998
R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
ng the biology of Mycobacterium tuberculosis from the complete genome
:: A70500; MUID:98295987; PMID:9634230
5
nary; nucleic acid sequence not shown; translation not shown
DNA
<COL>
5; GB:Z77162; GB:AL123456; NID:g3261606; PID:CAB00923.1; PID:e255036;
rce: strain H37Rv

larity   6.4%;   Score 93;   DB 2;   Length 310;
Conservative 24.4%; Pred. No. 4.8;   Indels 120;   Gaps   18;
             25;   Mismatches 103;

SLGSRDGGA-----VRQAQPPAPMAARRSQRRRG-------------------R 47
      :|: || |||||:|:|||| ||:||
SSGNRQISVAELLARQGVTGAP--ARRRRRRGDSDAITVAELTGEIPIIRDDHHH 63

GTALLVPLALGLGLALACLGLLLAVVSLGSRASLSAQEPAQEELVAEE-------- 99
||:|| | |: | ::|::|:|||||| |:||:|| ||||
AHASQSPAANGR-------------VQVGEAAPQSPAEPVAEQ-VAEEPTRTVYWS 109

SELNPQTEESQDPAPFLNRLVRPRR-----SAPKGRKTRARRAI---AAHY------ 146

Db   110 QPEPRWPKSPPQDRRESGPELSEYPRLRHTHSDRAPAGPPSGAEHMSPDPVBH'
            :: :||: | : : || ||| | :|:
Qy   147 ---------EVHPRPGQDG------AQAGVDGTVSGWEEARINSSS------

Db   170 DVLDTEVGEAEAETEVREAQPGRGERHAAAAAGTDVEGDGAAEARVARRALDV
                  |||||::|:| ||||
Qy   183 RQIGEFIVTR---------AGLYYLYCQVHFDE---GKAVYLKLDL---LVD(
               |:: ||||||||||| :|: |:|| |||
Db   230 ---GALVVLQSILAVAFGAGLF-----IAFDQLWRWNSIVALVLSVMVILGLVV!

Qy   226 CLEEFSAT-------AASSLGPQLRLCQ 246
               |||:| |:|:|||||
Db   282 KTEDIASTLIAVAVGALITLGP-LALLQ 308

RESULT 11
B34768
ORF5 protein - Orf virus (strain NZ2)
C;Species: Orf virus
C;Date: 23-Aug-1991 #sequence_revision 23-Aug-1991 #text_change 08-Oc
C;Accession: B34768
R;Fraser, K.M.; Hill, D.F.; Mercer, A.A.; Robinson, A.J.
Virology 176, 379-389, 1990
A;Title: Sequence analysis of the inverted terminal repetition in the
A;Reference number: A34768; MUID:90266454; PMID:2129563
A;Accession: B34768
A;Status: Preliminary
A;Molecule type: DNA
A;Residues: 1-351 <FRA>
A;Cross-references: GB:M30023; EMBL:M37623; NID:g3332561; PID:AAA4678

Query Match   6.4%;   Score 92;   DB 2;   Length 351;
Best Local Similarity 23.0%;   Pred. No. 6.6;
Matches   45; Conservative 18;   Mismatches 73;   Indels   60;

Qy   14 PLPRSLGSRDGGAVRQAQPPAPMAARRSQ-RRRGRRGEPGTALLVPLALGLGLAL
          |||| |:|||| | |||:
Db   211 PLPRRAAR---GRQRGQPPPRARRAQQPRRRAPRAAAAG---------------

Qy   73 LLAVVSLGSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQDPAPFLNRLVRPR
            :| :|||||| ||||
Db   248 --------------------------ARRGRGPAPRQQQRQPRVQRAAAQRRAQQR

Qy   133 GRKTRARRAIAAHYE-VHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYNRQIG
          |: ||||| |||| ||||
Db   284 PRRVRARRARARRQQRAHQR--RRGRRARTRCSTS-----RVVSKD----SREVG

Qy   192 RAGLYYLYCQVHFDEG 207
          :: :||:|||||||
Db   333 KERRYIRRVLLHFEDG 348

RESULT 12
T35203
probable two-component sensor - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 31-Jar
C;Accession: T35203
R;Seeger, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.J
submitted to the EMBL Data Library, April 1998
A;Reference number: Z21571
A;Accession: T35203
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-566 <SEE>
A;Cross-references: EMBL:AL022374; PID:CAA18527.1; GSPDB:GN00070; SCC
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SC5B8.19c
C;Superfamily: two-component sensor histidine kinase; sensor histidine

Query Match   6.4%;   Score 92;   DB 2;   Length 566;

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an c
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: D83231
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-664 <STO>
A;Cross-references: GB:AE004753; GB:AE004091; NID:g9949433; PIDN:AAG0E
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA3305

Query Match   6.3%;   Score 91.5; DB 2; Length 664;
Best Local Similarity   25.8%;   Pred. No. 15;
        Matches   77; Conservative   28; Mismatches 117; Indels   77; (

Qy   28 RQAQPPAPMAARR-------SQR------RRGRGREPGTALLVPLALGLGLLALACI
           ||| :   :: : |||| | | ||| ||||| :: :||| |
Db   194 RQRQPQGLLNALSKIVEVDAQRDHAWFEGERGRRRAGALALLSRDLLSL--

Qy   75 AVVSLGSRASLSAQEPAQEE-------LVAEEDQDPSELNPQTEESQDPAPFLNRL
           :|||||||||||||||||| |:|||||||||||||||||||||||||||
Db   246 ARGVARQRARLSEEERRVERWLAALALASALEGTDPASMQALREELAQVA-----

Qy   129 SAPKGRKTARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYNR
           ::| ||||||||||||| ||||||::
Db   300 SNDQ-RYLILTRCSVLLKAVN---------AEKGMRAVASGEVEGRVGSAGTLSWHR

Qy   187 EFIVTRAGL------YYLY------------------CQVHFDEGKAVYLKLDLLIV
           |||| ||| |:|| ||| | : :|||| || |
Db   352 LFYGTRSALALLGLSVYWIYTAWPAASGAMLLAAVVCSLFANRDNAVAIGLSFLR

Qy   224 LRCLEEFSATAASSLGPQLRLCQVSG--LLALRPGSSLRIRTLPWAHLKAAPFL1
           | ::   ::|||| ||| ||| ||:||||| |:||||||||||| |||
Db   412 I---------PAAMLVSQWLLPQWNGFPLLCLAMGVPLFFATLGMAVPVTAGTA1

RESULT 15
H83044
2,4-dienoyl-CoA reductase FadH2 PA4814 [imported] - Pseudomonas aeruc
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-D
C;Accession: H83044
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, I
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, I
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: H83044
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-681 <STO>
A;Cross-references: GB:AE004894; GB:AE004091; NID:g9951076; PIDN:AAG
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: fadH2; PA4814
C;Superfamily: Methylophilus methylotrophus W3A1 trimethylamine dehy

Query Match   6.3%;   Score 91; DB 2; Length 681;
Best Local Similarity   26.3%;   Pred. No. 17;
        Matches   86; Conservative   31; Mismatches 104; Indels   79;

Qy   6 FEISARRLPLPRSLGSRDGGAVRQAQPPAP----MAARRSQRRRGR---
           | :| ||| ||||| : ||||| |||||
Db   442 FRVRLERLGVDLRLGHR----VRQGELDGQPDDVVVATGIQPRRRIDGIGGP1

Qy   48 --RGEP-GTALLVPLALGLGLLALACLGLLLAVVSLGSRASLSAQEPAQEELVAI
             ||| |||||||||| ||||| || |||||||||||||||||||||||||
Db   498 VLRGAPVGARVAIVGAGGIGFDVA--AFLVAAPSDG------QPRALGEWLAI

```
PQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQDGAQAGVDGT 164
   ::||                ||          : :  |      |||
BQPGGLREPAP-----TRPARQV------------WLLQRKPGAPGAQL---GK     583

VEEARINSSSPLRYN--RQIGEFIVTRAGLYLYCQVHFDEGKAVYLKLD------LL 217
|   ::||:|:|      |:||    | :  ||  | : : ::|
VVH-----RAHLRHNAVRMLG------GVEYL----KIDE-RGLLIRVDGVERWLE  626

JL------ALRCLEEFSATAASSLGPQLRL-----CQVSGLL----ALRPGSSLRIR 262
|:  ||   ||  |   |        ||  | :|:| :|:|
JVICAGQEPLRELQIRQATES------LRFHLIGGARVAGELDAKRAIREGAMLAAR  680
```

April 7, 2004, 17:47:23
3 secs

```
        GenCore version 5.1.6
opyright (c) 1993 - 2004 Compugen Ltd.

n search, using sw model

il  7, 2004, 17:41:27 ; Search time 39.6149 Seconds
                        (without alignments)
                2261.954 Million cell updates/sec

09-245-198A-4
4                      .......PWAHLKAAPFLTYFGLFQVH 284
SLLDFEISARRLPLPRSLG

SUM62
op 10.0 , Gapext 0.5

7041 seqs, 315518202 residues

s satisfying chosen parameters:      1017041

th: 0
th: 2000000000

nimum Match 0%
ximum Match 100%
sting first 45 summaries

SUMMARIES
```

TREMBL_25:*
sp_archea:*
sp_bacteria:*
sp_fungi:*
sp_human:*
sp_invertebrate:*
sp_mammal:*
sp_mhc:*
sp_organelle:*
sp_phage:*
sp_plant:*
sp_rodent:*
sp_virus:*
sp_vertebrate:*
sp_unclassified:*
sp_rvirus:*
sp_bacteriap:*
sp_archeap:*

the number of results predicted by chance to have a
than or equal to the score of the result being printed,
d by analysis of the total score distribution.

| ry | | | | |
ch	Length	DB	ID	Description
.9	410	11	Q8BXS2	Q8bxs2 mus musculu
.4	330	4	Q8IZK7	Q8izk7 homo sapien
.0	261	5	Q8MRW2	Q8mrw2 drosophila
.0	325	5	Q9V5G2	Q9v5g2 drosophila
.0	409	5	Q8MY88	Q8my88 drosophila
.8	415	5	Q8MUJ1	Q8muj1 drosophila
.8	409	5	Q8IGD3	Q8igd3 drosophila
.2	398	6	Q8MK49	Q8mk49 sorex ciner
.2	532	16	Q82HP6	Q82hp6 streptomyce
.1	565	16	Q9KY66	Q9ky66 streptomyce
.0	955	10	Q84T85	Q84t85 oryza sativ
.0	967	10	Q7XLL4	Q7xll4 oryza sativ
.0	643	16	Q9KZ17	Q9kz17 streptomyce
.0	375	16	Q9RRH5	Q9rrh5 deinococcus
.0	893	16	Q8IZX0	Q8izx0 streptomyce
.9	611	11	Q8KOM8	Q8k0m8 mus musculu

17	100	17	850	6.9	Q9JJ15	Q9jj15 mus
18	100	11	850	6.9	Q80XI6	Q80xi6 mus
19	99.5	16	378	6.9	Q9AAB9	Q9aab9 caul
20	99	4	330	6.9	Q8N5L1	Q8n5l1 homo
21	99	4	614	6.9	Q7Z4K2	Q7z4k2 homo
22	99	4	915	6.9	Q7Z5I1	Q7z5i1 homo
23	98	16	206	6.8	Q9S2W5	Q9s2w5 stre
24	97.5	16	694	6.8	Q82FL1	Q82fl1 stre
25	97.5	4	1560	6.8	Q96JP2	Q96jp2 homo
26	97	10	408	6.7	Q85SI5	Q85si5 oryz
27	97	4	926	6.7	Q9NYA0	Q9nya0 homo
28	97	4	1058	6.7	Q9Y4G2	Q9y4g2 homo
29	96.5	11	975	6.7	Q8BWB1	Q8bwb1 mus
30	96	10	629	6.6	Q8S1A6	Q8s1a6 oryz
31	95.5	4	536	6.6	Q9HB96	Q9hb96 homo
32	95.5	16	655	6.6	Q9FBR7	Q9fbr7 stre
33	95	16	340	6.6	Q7WF13	Q7wf13 bord
34	95	2	390	6.6	Q8KW28	Q8kw28 ruege
35	95	16	810	6.6	Q82K60	Q82k60 stre
36	94.5	5	748	6.5	Q8T2Y0	Q8t2y0 trypa
37	94.5	11	1696	6.5	Q9WTR8	Q9wtr8 ratt
38	94	16	340	6.5	Q7W3N7	Q7w3n7 bord
39	94	16	340	6.5	Q7W0H0	Q7w0h0 bord
40	94	16	448	6.5	Q886A1	Q886a1 pseu
41	94	16	937	6.5	Q93JD1	Q93jd1 stre
42	94	10	1910	6.5	Q7XU19	Q7xu19 oryz
43	93.5	16	776	6.5	Q9RJ01	Q9rj01 stre
44	93.5	16	854	6.5	Q9F2P0	Q9f2p0 stre
45	93.5	10	1038	6.5	Q9AS09	Q9as09 oryz

ALIGNMENTS

RESULT 1
Q8BXS2
ID Q8BXS2;   PRELIMINARY;   PRT;   410 AA.
AC Q8BXS2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Tumor necrosis factor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Retina;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotatio
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK044387; BAC31897.1; -.
DR PIR; PT0714; PT0714.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005164; F:tumor necrosis factor receptor binding; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR006052; TNF_family.
DR InterPro; IPR008983; TNF_like.
DR SMART; SM00207; TNF; 2.
DR PROSITE; PS00251; TNF_1; 1.
DR PROSITE; PS50049; TNF_2; 2.
SQ SEQUENCE 410 AA; 45881 MW; 590A4B74C33FB8D4 CRC64; .

Query Match   64.9%; Score 937.5; DB 11; Length 410;
Best Local Similarity 79.1%; Pred. No. 1.3e-73;
Matches 193; Conservative 11; Mismatches 29; Indels 11; G

Qy   36 MAARSQRRRGRRGEPGTALLVPLALGLGLALACLGLLLAVVSLGSRASLSA-QEP
            ||||||||||||||| || || |||| ||||||||||||||| |||||| :||| |||

```
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   SD18286p.
GN   EIGER OR CG12919.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA   George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA   Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park
RA   Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M
RA   Celniker S.;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY119233; AAM51093.1; -.
DR   FlyBase; FBgn0033483; eiger.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005164; F:tumor necrosis factor receptor binding; IEA.
DR   GO; GO:0006955; P:immune response; IEA.
DR   InterPro; IPR006052; TNF_family.
DR   InterPro; IPR008983; TNF_like.
DR   SMART; SM00207; TNF; 1.
DR   PROSITE; PS002251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 1.
SQ   SEQUENCE  261 AA;  29780 MW;  13B6D5A04BC9122C CRC64;

Query Match     8.0%;  Score 116;  DB 5;  Length 261;
Best Local Similarity  25.4%;  Pred. No. 0.048;
Matches  57;  Conservative  32;  Mismatches  97;  Indels  38;

Qy   82 RASLSAQEPAQEELVAE----EDQDPSELNPQTEESQDPAPFLNR----------
                 ||:|        ::|| |:||||||||||||||||||
Db   54 RKSRSIADVRNEEQNIQGNHTELQEKSSNEATSKES--PAPLHRRRMHSRHRHL

Qy  128 RSAPKGRKTRARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYN
              ||||: :|| :|    | ||:|||||| | |||||||||
Db  112 ESLLSARSEDSRP--AAHFHLSSRRHQGSM-GYHGDMYIGNDNERNSYQG-HFQ

Qy  188 FIVTRAGLYLYCQV-----HFDEGKAVYLKLDLLVDGVLALRCLEEFSATAASS
              |:|  | |:|      ::|: |
Db  168 LTVTNTGLYYVVAQICYNNSHDQNGFIVP-----QGDTPFLQCLN----TVPTN

Qy  243 RLCQVSGLLALRPGGSLRIRTL---PWAHLKAAPFLTYGLFQV 283
              | : :||:          |::  ||:::|
Db  218 HTCHTSGLIHLERNERIHLKDIHNDRNAVLREGNNRSYFGIFKV 261


RESULT 4
Q9V5G2
ID   Q9V5G2
AC   Q9V5G2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   CG12919 protein.
GN   EIGER OR CG12919.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.
```

```
RRSQRRGRRGEPGTALLAPLVLSLGLALACLGLLLVVVSLGSWATLSAQQEPSQEE 60

EEDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQ 154
|: :| |||||||||||||||||||||||| ||||||| |||||||||||||||
EDRREPPELNPQTEESQDVVPFLEQLVRPRRSAPKGRKARPRRAIAAHYEVHPRPGQ 120

QAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYLYCQVHFDEGKAVYLKL 214
|||||||||||| :|||||||||:||||||| :||||||||||||||||||||||
QAGVDGTVSGWEETKINSSSPLRYDRQIGEFTVIRAGLYLYCQVHFDEGKAVYLKL 180

JDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALR------PGSSLRIRTLP 265
 :||||||||||||||||||| ||||||| :  :|
NGVLALRCLEEFSATAASSPGPQLRLCQTE-LQSLRREVSRLQRSGPSQKQGERP 239

_269

_243

PRELIMINARY;    PRT;   330 AA.

(TrEMBLrel. 23, Created)
(TrEMBLrel. 23, Last sequence update)
(TrEMBLrel. 25, Last annotation update)

(Human).

etazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
heria; Primates; Catarrhini; Hominidae; Homo.
506;

N.A.
9924; PubMed=12411489;
B., Medema J.P., Lopez-Fraga M., Lozano J.C.,
M., Picard A., Martinez-A C., Garcia-Sanz J.A.,

is hybrid mRNA encodes TWE-PRIL, a functional cell surface
fusion protein.";
?11-5720(2002).
51; AAL90443.1; -.
?0; C:membrane; IEA.
?4; F:tumor necrosis factor receptor binding; IEA.
?5; P:immune response; IEA.
?006052; TNF_family.
?008983; TNF_like.
?; TNF; 1.
?7; TNF; 1.
?251; TNF_1; 1.
?049; TNF_2; 2.
?0 AA;  36588 MW;  FC6F3BCA29C029AE CRC64;

larity   58.4%;  Score 844;  DB 4;  Length 330;
Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

RRSQRRGRRGEPGTALLVPLVLALGLGLALACLGLLLAVVSLGSRASLSAQEPAQEL 95
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
RRSQRRGRRGEPGTALLVPLVLALGLGLALACLGLLLAVVSLGSRASLSAQEPAQEL 60

EDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQD 155
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
EDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQD 120

AGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQ 201
|||||||||||||||||||||||||||||||||||||||||||
AGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQ 166

PRELIMINARY;    PRT;   261 AA.
```

```
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   TNF superfamily ligand, Eiger (Tumor necrosis factor family member
DE   DT1).
GN   EIGER OR DT1 OR CG12919.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22060500; PubMed=12065414;
RA   Igaki T., Kanda H., Yamamoto-Goto Y., Kanuka H., Kuranaga E.,
RA   Aigaki T., Miura M.;
RT   "Eiger, a TNF superfamily ligand that triggers the Drosophila JNK
RT   pathway.";
RL   EMBO J. 21:3009-3018(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Inohara N., Nunez G.;
RT   "DT1, a Drosophila tumor necrosis factor family member.";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB073865; BAC00950.1; -.
DR   EMBL; AF149799; AAO15310.1; -.
DR   FlyBase; FBgn0033483; eiger.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005164; F:tumor necrosis factor receptor binding; IEA.
DR   GO; GO:0006955; P:immune response; IEA.
DR   InterPro; IPR006052; TNF_family.
DR   InterPro; IPR008983; TNF_like.
DR   SMART; SM00207; TNF; 1.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 1.
SQ   SEQUENCE   409 AA;  46331 MW;  8306AECCE14397B8 CRC64;

Query Match             8.0%;  Score 116;  DB 5;  Length 409;
Best Local Similarity   25.5%;  Pred. No. 0.084;
Matches   56;  Conservative  33;  Mismatches  95;  Indels  36;  Ga

Qy      82 RASLSAQEPAQEELVAE----EDQDPSELNPQTEBSQDPAPFLNRLVRPRRSAPKG
           ||::  | ::|   : |     ||||| : ||:       :||||:    ||:::|
Db     208 RKSRSIADVRNEEQNIQGNHTELQEKSSNEATSKES--PAPLHHR----RRMHSRH

Qy     138 ARRA----IAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYNRQIGE
           |::|    || |:  |  : |: :::|  |:|||||||||||||||:||:
Db     262 VRKARSEDSRPAAHFHLSSRRRHQSM-GYHGDMYIGNDNERNSYQG-HFQTRDGV.

Qy     192 RAGLYLYCQV-----HFDEGKAVVLKLDLLVDGVLALRCLEEFSATAASSLGPQL
           :||| | :|:     ||::|:||:||||||||||||:||||||||||||||||||
Db     320 NTGLIYYVYAQICYNNSHDQNGFIVP-----QGDTPFLQCLN----TVPTNMPHKV

Qy     247 VSGLIALRPGSSLRIRTL--PWAHLKAAPPLTYFGLFQV 283
           .||| | | :|::|:| |  |||| |||||||||:|:.|
Db     370 TSGLIHLERNERIHLKDIHNDRNAVLREGNNRSYFGIFKV 409
```

```
RESULT 6
Q8MUJ1
ID   Q8MUJ1      PRELIMINARY;      PRT;   415 AA.
AC   Q8MUJ1;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Eiger (DARTH).
GN   EIGER OR CG12919 OR DARTH.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
```

Left column (partially cut off):

```
ortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
a C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
oayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
asu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
enos P.V., Berman B.P., Bhandari D., Bolshakov S.,
tchan M.R., Bouck J., Brokstein P., Brottier P.,
usam D.A., Butler H., Cadieu E., Center A., Chandra I.,
awley S., Dahlke C., Davenport L.B., Davies P.,
Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
p L.E. Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
vangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
rielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
g F., Gorrell J.H., Gu Z., Guan P., Harris M.,
arvey D., Heiman T.J., Hernandez J.R., Houck J.,
ston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
ush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
odira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
B., McIntosh T.C., McLeod M.P., McPherson D.,
ilshina N.V., Mobarry C., Morris J., Moshrefi A.,
y M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
elson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
mington K., Saunders R.D.C., Scheeler F., Shen H.,
en-Kiamos I., Simpson M., Skupski M.P., Smith T.,
dling A.C., Stapleton M., Strong R., Sun E.,
ector C., Turner R., Venter E., Wang A.H., Wang X.,
ssarman D.A., Weinstock G.M., Weissenbach J.,
Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
ong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
ers E.W., Rubin G.M., Venter J.C.;
quence of Drosophila melanogaster.";
85-2195(2000).
; AAF58848.1; -.
033483; eiger.
; C:membrane; IEA.
; F:tumor necrosis factor receptor binding; IEA.
; P:immune response; IEA.
06052; TNF_family.
08983; TNF_like.
; TNF; 1.
51; TNF_1; 1.
49; TNF_2; 1.
, AA;  36862 MW;  6E5CCB69694F1A3A CRC64;

rity           8.0%;  Score 116;  DB 5;  Length 325;
               Pred. No. 0.063;
onservative  32;  Mismatches  97;  Indels  38;  Gaps  10;

AQEPAQEELVAE----EDQDPSELNPQTEBSQDPAPFLNR-------LVRPR 127
IADVRNEEQNIQGNHTELQEKSSNEATSKES--PAPLHHRRRMHSRHRHLLVRKG 175

GRKTRARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYNRQIGE 187
SARSEDSRP--AAHFHLSSRRRHQSM-GYHGDMYIGNDNERNSYQG-HFQTRDGV 231

RAGLYLYCQV-----HFDEGKAVVLKLDLLVDGVLALRCLEEFSATAASSLGPQL 242
NTGLIYYVYAQICYNNSHDQNGFIVP-----QGDTPFLQCLN----TVPTNMPHKV 281

VSGLIALRPGSSLRIRTL--PWAHLKAAPPLTYFGLFQV 283
TSGLIHLERNERIHLKDIHNDRNAVLREGNNRSYGIFKV 325

PRELIMINARY;      PRT;   409 AA.
(TrEMBLrel. 22, Created)
```

DR   FlyBase; FBgn0064801; BcDNA:RH51659.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005164; F:tumor necrosis factor receptor binding; IEA.
DR   GO; GO:0006955; P:immune response; IEA.
DR   InterPro; IPR006052; TNF_family.
DR   InterPro; IPR008983; TNF_like.
DR   SMART; SM00207; TNF; 1.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 1.
SQ   SEQUENCE   409 AA;   46401 MW;   FC2B9BD9E012D257 CRC64;

Query Match
Best Local Similarity   7.8%; Score 113; DB 5; Length 409;
Matches   54; Conservative   36; Mismatches   94; Indels   44; G

Qy   84 SLSAQEPAQEL-----VAEDQDPSELNPQTEE--------SQDPAPFLNRLVR
Db 198 SYNAHKKQERKSRSIADVRNEEQNIQGNHTELQEKSSNEATSKERPAPLHHR---

Qy 130 APKGRKTRARRA-----IAAHYEVHRPPGQDGAQAGVDGTVSGWEEARINSSSPL
Db 254 HSRHRHLLVRKARSEDSRPAAHFHLSSRRRHQGSM-GYHGDMYIGNDNERMSYQG-

Qy 184 QIGEFIVTRAGLYLYCQV-----HFDEGKAVYLKLDLLVDGVLALRCLEEFSATA
Db 312 RDGVLTVTNTGLYYVVAQICYNNSHDQNGFIVF-----QGDTPFLQCLN----TV

Qy 239 GPQLRLCQVSGLLALRPGSSLRIRTL---PWAHLKAAPFLTYFGLFQV 283
Db 362 PHKVHTCHTSGLIHLERNERIIHLKDIHNDRNAVLREGNNRSYFGIFKV 409

RESULT 8
Q8MK49
ID   Q8MK49   PRELIMINARY;   PRT;   398 AA.
AC   Q8MK49;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Alpha 2B adrenergic receptor (Fragment).
GN   ADRA2B.
OS   Sorex cinereus (Masked shrew).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Insectivora; Soricidae; Soricinae; Sorex.
OX   NCBI_TaxID=36803;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21608557; PubMed=11743200;
RA   Murphy W.J., Eizirik E., O'Brien S.J., Madsen O., Scally M.,
RA   Douady C.J., Teeling E., Ryder O.A., Stanhope M.J., de Jong W.W.,
RA   Springer M.S.;
RT   "Resolution of the early placental mammal radiation using Bayesia
RT   phylogenetics.";
RL   Science 294:2348-2351(2001).
DR   EMBL; AJ315936; CAC87000.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin..
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER   1   1
FT   NON_TER   398   398
SQ   SEQUENCE   398 AA;   43576 MW;   D57E67B689535E27 CRC64;

Query Match
Best Local Similarity   7.2%; Score 104; DB 6; Length 398;
Matches   55; Conservative   23; Mismatches   76; Indels   64; C

923; PubMed=12176339;
n M., Basler K.;
TNF Signaling Mechanisms, JNK-Dependent Apoptosis
Eiger, the Drosophila Homolog of the TNF Superfamily.";
2:1263-1268(2002).

N.A.
938; PubMed=12894227;
Maaty W.S., Chen P., Tomar R.S., Eby M.T., Chapo J.,
ore N., Zachariah S., Sinha S.K., Abrams J.M.,
.;
s receptor, Wengen, comprise a TNF-like system in

860-4867(2003).
6; AAM76710.1; -.
1; AAM66763.1; -.
0033483; eiger.
0; C:membrane; IEA.
4; F:tumor necrosis factor receptor binding; IEA.
5; P:immune response; IEA.
006052; TNF_family.
008983; TNF_like.
7; TNF; 1.
251; TNF_1; 1.
049; TNF_2; 1.
5 AA;   46918 MW;   E087A26DE222D2BF CRC64;

arity   8.0%; Score 116; DB 5; Length 415;
onservative   32; Mismatches   97; Indels   38; Gaps   10;

SAQEPAQEELVAE---EDQDPSELNPQTEESQDPAPFLNR------------LVRPR 127
SIADVRNEEQNIQGNHTELQEKSSNEATSKES--PAPLHHRRRMHSRHRHLLVRKG 265

KGRKTRARAIAAHYEVHPRPGDGAQAGVDGTVSGWEEARINSSSP1RYNRQIGE 187
SARSEDSRP--AAHFHLSSRRRHQGSM-GYHGDMYIGNDNERNSYQG-HFQTRDGV 321

RAGLYYLYCQV-----HFDEGKAVYLKLDLLVDGVLALRCLEEFSATAASSLGPQL 242
NTGLYYVYAQICYNNSHDQNGFIVF------QGDTPFLQCLN---TVPTNMPHKV 371

VSGLLALRPGSSLRIRTL---PWAHLKAAPFLTYFGLFQV 283
TSGLIHLERNERIIHLKDIHNDRNAVLREGNNRSYFGIFKV 415

RELIMINARY;   PRT;   409 AA.

TrEMBLrel. 23, Created)
TrEMBLrel. 23, Last sequence update)
TrEMBLrel. 25, Last annotation update)

lanogaster (Fruit fly).
tazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
opterygota; Diptera; Brachycera; Muscomorpha;
Drosophilidae; Drosophila.
27;
1 N.A.

Brokstein P., Hong L., Agbayani A., Carlson J.,
avez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
nzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
ungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
uuanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
V-2002) to the EMBL/GenBank/DDBJ databases.
8; AAN71595.1; -.

```
Qy    159  ------------AGVDGTV-SGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYI
                                   :  ||  |:: :::
Db    428  TGIDFPVSYGTTVLAATDGTVRTQWNSAYGNMA----------IVTAKDGTETWVI

Qy    204  FDEGKAVYLKLDLLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLI
                   |||:  | :| ::|          ::||
Db    475  LSTYKVASGTTVKAGDPI-------AFSGNSGNSTGPHLH--------FEVRPAGGSI

Qy    264  LPW  266
           |||
Db    521  LPW  523
```

RESULT 10

```
Q9KY66
ID   Q9KY66          PRELIMINARY;         PRT;     565 AA.
AC   Q9KY66;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative peptidase.
GN   SC04798 OR SCD63A.09C.
OS   Streptomyces coelicolor.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3 (2);
RA   Brown S.F., Harris D.;
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3 (2);
RA   Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3 (2);
RX   MEDLINE=97000351; PubMed=8843436;
RA   Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
RA   Kinashi H., Hopwood D.A.;
RT   "A set of ordered cosmids and a detailed genetic and physical map
RT   the 8 Mb Streptomyces coelicolor A3 (2) chromosome.";
RL   Mol. Microbiol. 21:77-96(1996).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3 (2) / M145;
RX   MEDLINE=2199610; PubMed=12000953;
RA   Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA   Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA   Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M
RA   Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S
RA   Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.
RA   Rabbinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA   Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor ]
RA   Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA   Hopwood D.A.;
RT   "Complete genome sequence of the model actinomycete Streptomyces
RT   coelicolor A3(2).";
RL   Nature 417:141-147(2002).
DR   EMBL; AL939121; CAB92661.1; -.
DR   GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR002886; Peptidase_M37.
DR   Pfam; PF01551; Peptidase_M37; 1.
KW   Complete proteome.
SQ   SEQUENCE   565 AA;  58070 MW;  7D0418D480C6A284 CRC64;
```

Query Match   7.1%;   Score 103;   DB 16;   Length 565;
Best Local Similarity   21.2%;   Pred. No. 1.7;
Matches   70;   Conservative   33;   Mismatches   99;   Indels   128;   G.

(left column, rotated)

```
SRDGGAVRQA-QPP--APMAAR--------------------------RSQRRRGRRGEP  51
 :   || |  | :::|
ASGKGESQAGQAPRGAFVSAKLPSLASPLVVAGDAPGLSKPPRENMETGAPEP  253

VPLALGLGLAL-------ACLGLLLAVVSLGSRASLSAQEPAQEELVABEDQDPSE  105
                    |:|| |:|       ||:: ||
PPSWPALPSAAQDLKEGACV----------------VSABEEAEEDYAEEEEEEE  299

EESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHRPPGQDGAQAGVDGTV  165
 ||       | :|||    ::|
L----PAPTIAL------SPALRQPQGARVLAT-----LRGQVLLSRAVGTAV  343

ARINSSSPLRYN---RQIGEFIVTRAGLYYLY  199
 ||||||| |   ||| : ||
RRAQLSREKRFTFVLAVVIGVFVLCWFPFFSY  381
```

PELIMINARY;   PRT;   532 AA.

TrEMBLrel. 24, Created)
TrEMBLrel. 24, Last sequence update)
TrEMBLrel. 25, Last annotation update)
dase.

vermitilis.
nobacteria; Actinobacteridae; Actinomycetales;
e; Streptomycetaceae; Streptomyces.
03;

N.A.
/ ATCC 31267 / NCIMB 12804 / NRRL 8165;
03; PubMed=11572948;
la H., Ishikawa J., Hanamoto A., Takahashi C.,
ukahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
iba T., Sakaki Y., Hattori M.;
ce of an industrial microorganism Streptomyces
ducing the ability of producing secondary

ad. Sci. U.S.A. 98:12215-12220(2001).

N.A.
/ ATCC 31267 / NCIMB 12804 / NRRL 8165;
06; PubMed=12692562;
kawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
tori M., Omura S.;
me sequence and comparative analysis of the industrial
Streptomyces avermitilis.";
l. 21:526-531(2003).
; BAC71174.1; -.
; F:metalloendopeptidase activity; IEA.
; P:proteolysis and peptidolysis; IEA.
02886; Peptidase_M37.
Peptidase_M37; 1.
ome.
2 AA;   55998 MW;   6E3F1CEC61E5A738 CRC64;
```

arity   23.4%;   Pred. No 1.4;   DB 16;   Length 532;
onservative   22;   Mismatches   107;   Indels   103;   Gaps   14;

```
BLGSRSDGGAVRQAQPPAPMAARRSQRRRGRGEPGTALLVPLALGLALACLGLL  73
 |||
--GSR-------AASRTASRSRRIPARRSALLTVAVPSACVMGVA----GIA  316

BLGSRASLSAQEPAQ--------------------BELVABED--------QDPS  104
 ||||                              :||:
SVGSSDSKDAHTTASAADAIAVKPSAANNKLDTQIESLSAGADDFADRASRTQERI  376

PQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHRPPGQDGAQ-----  158
 :| |  :||   |:|:|   :|:|
QQAAEKQKAAA--EAARKERLRPKFALPVAQKGLSAYY------GQAGINWMSVH  427
```

Db   233 TIEAIRAKRQQLQQPRHAAP 252
                  :|:||:||

RESULT 12
Q7XLL4
ID   Q7XLL4   PRELIMINARY;   PRT;   967 AA.
AC   Q7XLL4;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   OSJNBa0094P09.13 protein.
GN   OSJNBa0094P09.13.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,
RA   Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,
RA   Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li
RA   Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.
RA   Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
RA   Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin
RA   Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,
RA   Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL731625; CAE05074.1; -.
SQ   SEQUENCE   967 AA; 106012 MW; 3FA9D0CFE245B970 CRC64;

Query Match   7.0%;   Score 101.5;   DB 10;   Length 967;
Best Local Similarity 23.9%;   Pred. No.4.6;
Matches   70; Conservative   35; Mismatches 123; Indels   65; C

Qy   17 RSLGS--------RDGGAVRQ---------AQPPAPMAARRSQRRRGRGEPGTALLVE
             |||| ::| ||||||
Db   352 RKLGGTPPPSPPRGGGAVRASSRRPEGAAPTSQPEGERKKKRLRKTGE------------

Qy   63 LGLALACLGLLLA---------------VVSLGSRASLSAQEPAQEELVAEE----DQI
             |||: |:|: |||
Db   400 ---TEPCRGNLISPPRWSFNRPPRSDVPSHPSRHPKSSQSEARDPAAAAERRRESI

Qy   107 NPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGGQDGAQAGVI
             |||:|| :|||| |||: |||||||
Db   457 ADRLRAEAARRAAR-VRQAEEAAREEAAARQAEEAAKEEADRARQAEATA------

Qy   167 GWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYLKLDLLVD----
             :||:   :|||: :|||:   :|:|:   :||:|
Db   511 GAPWPRANTAVLDGFNAQVEALRAERAELEAAWTRV--EEGRR---SVDAMVEVGH

Qy   220 -GVLALRCLEEFSATAASLGPQLRLCQVSGLLALRPGSSLRIRTLPW-AHLK 2;
             :|||: ||:: :|||: ||:: ||:|
Db   566 RHVSELEARKAALAEIAREVEEERAAALISTTAMVEAQDTLRLQHASWEAELK 61

RESULT 13
Q9KZ17
ID   Q9KZ17   PRELIMINARY;   PRT;   643 AA.
AC   Q9KZ17;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein SCO2220.
GN   SCO2220 OR SC10B7.15.
OS   Streptomyces coelicolor.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.

DR   InterPro; IPR007863; Peptidase_M16_C.
DR   Pfam; PF05193; Peptidase_M16_C; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   375 AA; 39596 MW; 197B397658BD30B3 CRC64;

Query Match
Best Local Similarity   7.0%;   Score 100.5;   DB 16;   Length 375;
Matches   69;   Conservative   27;   Mismatches   84;   Indels   107;   Ga

Qy   25 GAVRQAQPPAPMAA-----RRSQRRGRRGBEPGT--ALLVPLALGLGLALACLGLLI
              ||| ||:||::||| ||| ||||
Db   5 GRGHRCQPTPGRAATARRLRRPGRAPGRRGQPGSHPAARQRLRADLGAALA------

Qy   78 SLGSRASLSAQE-PAQEELVAEEDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPK(
         :||::|||:|| || |: |:
Db   59 DVVTRPTLPAGELPTLLDLAR--------QDLESLDDSP---------

Qy 137 RARRAIAAHYEVHPR-PGQDGAQAG--VDGTVSGWEEARINSSSPLRYNRQIGE---
                   ||:|| ||| ::|::||||
Db   90 TDRLAVQARAATFPRLPGDPGAGFGHPISGTLAGLE-RLSAASLRAHWARFGQRGS

Qy 189 IVTRAGLIYLY--------------------CQVHFDEGKAVYLKLDLLVDGVLALRCLI
              ::|| :|:
Db 148 VVADADAQEVYELVAGLFADWQPGEDRMPAHFQPGLRLHL

Qy 232 ATAASSLGPQLRLCQVS----------GLLALRPGSSLRIRT 263
         :||| ||| ||||||||||:
Db 189 ---PSATGEQTHLSLVAPGPGPRDPDWLPWQLALTALSGGSASRLFT 232

RESULT 15
ID   Q81ZX0   PRELIMINARY;   PRT;   893 AA.
AC   Q81ZX0;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Putative serine/threonine protein kinase.
GN   PKN19 OR SAV5006.
OS   Streptomyces avermitilis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=33903;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX   MEDLINE=21477403; PubMed=11572948;
RA   Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA   Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA   Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT   "Genome sequence of an industrial microorganism Streptomyces
RT   avermitilis: deducing the ability of producing secondary
RT   metabolites.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX   MEDLINE=22608306; PubMed=12692562;
RA   Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba
RA   Sakaki Y., Hattori M., Omura S.;
RT   "Complete genome sequence and comparative analysis of the industri
RT   microorganism Streptomyces avermitilis.";
RL   Nat. Biotechnol. 21:526-531(2003).
DR   EMBL; AP005041; BAC72718.1; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_Thr_pkinase.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 2.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

M145;
10; PubMed=12000953;
Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
James K.D., Harris D.E., Quail M.A., Kieser H.,
eman A., Brown S., Chandra G., Chen C.W., Collins M.,
ser T., Larke L., Murphy L., Hornsby T., Howarth S.,
ieser T., Rajandream M.A., Rutherford K., Rutter S.,
nders D., Sharp S., Squares R., Squares S., Taylor K.,
tzorrek A., Woodward J., Barrell B.G., Parkhill J.,

me sequence of the model actinomycete Streptomyces
2).";
-147(2002).
; CAB90868.1; -.
rotein; Complete proteome.
AA; 66537 MW; 1217A5C88B3CD419 CRC64;

rity   7.0%;   Score 101;   DB 16;   Length 643;
nservative   28;   Mismatches   122;   Indels   72;   Gaps   10;

---LPRSLGSRD-GGAVRQAQ-----------PPAPMAARRSQRRRGRRGE 50
   :||||||| |||||| ||||||||||:|||||||||
GVRQVLRAAGVADPDAALREADGVPGQYGLLGSPEFDPCSLQARPTDLLRRROH- 223

LVPLALGLGLALACLGLLAVVSLGSRASLSAQEPAQEELVAEEDQDPSELNPQT 110
|||||||||||||||:|| ::|||:||:|| || |:
-TKAALVAGAALVVCGALLGLPGDGWGPDGAAAPPYAQNPAAEAALDPGRLTKAA 277

PAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEE 170
  :||| || |||:| |||:|||||
TSARTDFSVWPARGGLTGDEELLRRALA-----VWARPGESVGVSATPGTATG--- 330

SSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYLKLDLLVDGVLALRCLEEF 230
|||||||||| |:||||||||:|| |||:
-GP--------AGPPQLLYAGEVDTARVV-----VLHDGLRLVRYAEPK 365

SSL-------GPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKAA 272
:::   ::|||||||||||||
AALDFARTDGAGRAAATAVVLG----RADGNVRYLTAPWVTKAAA 411

ELIMINARY;   PRT;   375 AA.

rEMBLrel. 13, Created)
rEMBLrel. 13, Last sequence update)
rEMBLrel. 24, Last annotation update)
rotein DR2516.

ococcus-Thermus; Deinococci; Deinococcales;
; Deinococcus.
9;

N.A.
CC 13939 / DSM 20539 / NCIB 9279;
96; PubMed=10567266;
n J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
aft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
in H., Jiang L., Pamphile W., Crosby R., Shen M.,
, Lam P., McDonald L., Utterback T., Zalewski C.,
Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,

diodurans.
ce of the radioresistant bacterium Deinococcus

.";
71-1577(1999).
; AAF12062.1; -.
75264.
-

e/threonine-protein kinase; Complete proteome.
3 AA; 92763 MW; CC4DA95AFAFE2407 CRC64;

```
arity     7.0%;  Score 100.5;  DB 16;  Length 893;
          22.7%;  Pred. No. 5.1;
onservative  29;  Mismatches  83;  Indels  45;  Gaps     7;

RLPLPRSLGSRDGGAVRQAQPPAPMAARRSQRRGRRGEPGTALLVPLALGLGLAL  67
|||:|:|       |||:      : |:|:| ||      |        ||||||
RLPIVR---RRGELVRRRARLPATTPHARHKRAKDGRPPR----PKRLGRNLVL  660

:LLLAVVSL----------GSRA------SLSAQEPAQELVAEEDQDPSELN  107
:|:|:|:            ||::     :   |||:|   ||| : :
VLAAAVAYAMVFMPKASPNSDSGSRSDGDRTGAAGGVSPAPEQSGGGSEPRPDQTS  720

:ESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQDGAQAGVDGTVSG  167
:  |: |:   |    ::  :: : :|    :: :   :|   :
KS-----------PSGSGSTREQTTDPDVAQGFTLRKDP--EGFRVAV---ATG  763

RINSSSPLRYNRQIGEFTV  190
:  :|:      ||  |
PKNGRGRIVYSHGDFELIV  786
```

April 7, 2004, 17:46:45
secs

```
          GenCore version 5.1.6
opyright (c) 1993 - 2004 Compugen Ltd.

a search, using sw model

il  7, 2004, 17:38:07 ;  Search time 11.7171 Seconds
                          (without alignments)
                          1262.081 Million cell updates/sec

09-245-198A-4
4

SLLDFEISARRLPLPRSLG.........PWAHLKAAPFLTYFGLFQVH 284

SUM62
op 10.0 , Gapext 0.5

681 seqs, 52070155 residues

s satisfying chosen parameters:          141681

th: 0
th: 2000000000

nimum Match 0%
ximum Match 100%
sting first 45 summaries

issProt_42:*

the number of results predicted by chance to have a
than or equal to the score of the result being printed,
d by analysis of the total score distribution.

                        SUMMARIES
```

ry Length	DB	ID	Description
.8 249	1	TN12_HUMAN	O43508 homo sapien
.6 225	1	TN12_MOUSE	O54907 mus musculu
.6 272	1	TNF5_CHICK	Q9i8d8 gallus gall
.5 260	1	TNF5_CANFA	O97626 canis famil
.4 254	1	TNF9_HUMAN	P41273 homo sapien
.9 952	1	HDA7_HUMAN	Q8wui4 homo sapien
.7 441	1	CG22_ANTMA	P34801 antirrhinum
.6 201	1	TNFB_MACEU	Q9xt48 macropus eu
.5 280	1	TNF6_MACMU	Q9myl6 macaca mula
.5 690	1	RHO_MICLU	P52154 micrococcus
.5 310	1	Y497_MYCTU	Q11162 mycobacteri
.4 240	1	TN14_HUMAN	O43557 homo sapien
.4 280	1	TNF6_CERTO	Q9bdn1 cercocebus
.4 902	1	NFC4_HUMAN	Q14934 homo sapien
.3 760	1	MLH1_MOUSE	Q9jk91 mus musculu
.3 814	1	CADF_HUMAN	P55291 homo sapien
.2 707	1	JIP1_MOUSE	Q9wvi9 mus musculu
.2 280	1	MDCB_KLEPN	P71422 klebsiella
.2 316	1	TN11_MOUSE	O35235 m tumor nec
.1 278	1	TNF6_RAT	P36940 rattus norv
.1 281	1	TNF6_HUMAN	P48023 homo sapien
.1 422	1	GFI1_HUMAN	Q99684 homo sapien
.1 574	1	SEN3_HUMAN	Q9h414 homo sapien
5.0 197	1	TNFB_RABIT	P10154 oryctolagus
5.0 204	1	TNFB_BOVIN	Q06600 bos taurus
5.0 291	1	TN10_MOUSE	P50592 mus musculu
5.0 250	1	TNFC_MACEU	Q9xt47 macropus eu
5.0 139	1	YQFB_BACSU	P54467 bacillus su
5.0 205	1	TNFB_MARMO	Q9jm09 marmota mon
5.0 241	1	TN13_MOUSE	Q9d777 mus musculu
5.0 777	1	METE_CAUCR	Q9aaw1 caulobacter
5.0 933	1	VGLB_HSVA1	Q04463 herpesvirus
5.9 928	1	VGLB_HSVBP	P17471 bovine herp

34	85.5	5.9	932	1	VGLB_HSVBC	P12640 bovir
35	85	5.9	372	1	LMXB_MOUSE	O88609 mus t
36	85	5.9	379	1	LMXB_HUMAN	O60663 homo
37	84.5	5.9	401	1	AROC_MYCTU	P95013 mycol
38	84	5.8	284	1	TLX2_HUMAN	O43763 homo
39	84	5.8	310	1	TNFC_MARMO	Q9jm10 marmc
40	84	5.8	575	1	MIS_PIG	P79295 sus s
41	84	5.8	703	1	ZM15_HUMAN	Q9h091 homo
42	84	5.8	825	1	ICP0_HSV2H	P28284 herpe
43	83.5	5.8	416	1	RAGE_BOVIN	Q28173 bos t
44	83.5	5.8	505	1	TUB_MOUSE	P50586 mus n
45	83.5	5.8	545	1	RTN2_HUMAN	O75298 homo

ALIGNMENTS

```
RESULT 1
TN12_HUMAN
ID   TN12_HUMAN     STANDARD;      PRT;      249 AA.
AC   O43508; Q8WUZ7;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Tumor necrosis factor ligand superfamily member 12 (TNF-related we
DE   inducer of apoptosis) (TWEAK) (APO3 ligand).
GN   TNFSF12 OR APO3L OR DR3LG.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND N-TERMINUS OF SOLUBLE FORM.
RC   TISSUE=Fetal liver, and Tonsil;
RX   MEDLINE=98070415; PubMed=9405449;
RA   Chicheportiche Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H.,
RA   Hession C., Garcia I., Browning J.L.;
RT   "TWEAK, a new secreted ligand in the tumor necrosis factor family
RT   weakly induces apoptosis.";
RL   J. Biol. Chem. 272:32401-32410(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal kidney;
RX   MEDLINE=98228355; PubMed=9560343;
RA   Marsters S.A., Sheridan J.P., Pitti R.M., Brush J., Goddard A.,
RA   Ashkenazi A.;
RT   "Identification of a ligand for the death-domain-containing recept
RT   Apo3.";
RL   Curr. Biol. 8:525-528(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Tonsil;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
```

```
Qy  156 GAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYIYLYCQVHFDEGKAVY
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYIYLYCQVHFDEGKAVY

Qy  216 LLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKA
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 LLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKA

Qy  276 TYFGLFQVH 284
            |||||||||
Db  241 TYFGLFQVH 249
```

RESULT 2
TN12_MOUSE
```
ID   TN12_MOUSE     STANDARD;     PRT;     225 AA.
AC   O54907; Q9CTP2;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Tumor necrosis factor ligand superfamily member 12 (TNF-related w
DE   inducer of apoptosis) (TWEAK) (Fragment).
GN   TNFSF12.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Peritoneal macrophage;
RX   MEDLINE=98070415; PubMed=9405449;
RA   Chicheportiche Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H.,
RA   Hession C., Garcia I., Browning J.L.;
RT   "TWEAK, a new secreted ligand in the tumor necrosis factor family
RT   weakly induces apoptosis.";
RL   J. Biol. Chem. 272:32401-32410(1997).
RN   [2]
RP   SEQUENCE OF 83-225 FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Retina;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbus
RA   Schriml L.M., Stabili F., Suzuki R., Tomita M., Wagner L., Washi
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilmi
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
CC   -!- FUNCTION: Binds to FN14 and possibly also to TNRFSF12/APO3.
CC       inducer of apoptosis in some cell types. Promotes angiogenes
CC       the proliferation of endothelial cells. Mediates NF-KappaB
CC       activation (By similarity).
CC   -!- SUBUNIT: Homotrimer (Potential).
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein and secreted
CC       similarity).
CC   -!- TISSUE SPECIFICITY: Widely expressed.
CC   -!- PTM: The soluble form is produced from the membrane form by
CC       proteolytic processing (By similarity).
CC   -!- SIMILARITY: Belongs to the tumor necrosis factor family.
CC   -----------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a col
```

```
061; PubMed=10085077;
ang Y.C., Lund J.K., Chen Y.-W., Leal J.A., Wiley S.R.;
s angiogenesis and proliferation of endothelial cells.";
..274:8455-8459(1999).
    Binds to FN14 and possibly also to TNRFSF12/APO3. Weak
f apoptosis in some cell types. Mediates NF-KappaB
n. May promote angiogenesis and the proliferation of
al cells.
Homotrimer (Potential).
AR LOCATION: Type II membrane protein and secreted.
ECIFICITY: Highly expressed in adult heart, pancreas,
muscle, brain, colon, small intestine, lung, ovary,
spleen, lymph node, appendix and peripheral blood
es. Low expression in kidney, testis, liver, placenta,
d bone marrow. Also detected in fetal kidney, liver,
brain.
soluble form derives from the membrane form
lytic processing.
Y: Belongs to the tumor necrosis factor family.
Ref.3 sequence differs from that shown due to a
t in position 125.
----------------------------------------------------------
OT entry is copyright. It is produced through a collaboration
Swiss Institute of Bioinformatics and the EMBL outstation -
Bioinformatics Institute. There are no restrictions on its
profit institutions as long as its content is in no way
this statement is not removed. Usage by and for commercial
ires a license agreement (See http://www.isb-sib.ch/announce/
ail to license@isb-sib.ch).
----------------------------------------------------------
9; AAC51923.1; -.
2; AAC39724.1; -.
7; AAH19047.1; ALT_FRAME.
1927; TNFSF12.
-; -.
7;  C:integral to plasma membrane; TAS.
2; F:receptor binding; TAS.
7; P:induction of apoptosis; TAS.
5; P:signal transduction; TAS.
006052; TNF family.
009983; TNF_like.
7; TNF; 1.
7; TNF; 1.
251; TNF_1; FALSE_NEG.
049; TNF_2; 1.
iogenesis; Apoptosis; Transmembrane; Glycoprotein;
    1      249    TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
                  MEMBER 12, MEMBRANE FORM.
   94      249    TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
                  MEMBER 12, SECRETED FORM.
    1       21    CYTOPLASMIC (POTENTIAL).
   22       42    SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
                  (POTENTIAL).
   43      249    EXTRACELLULAR (POTENTIAL).
   93       94    CLEAVAGE.
   39      139    N-LINKED (GLCNAC...).
9 AA;   27216 MW;  E660843361C28EBA CRC64;
larity   87.8%;   Score 1268;  DB 1;  Length 249;
    100.0%;   Pred. No. 4e-92;
Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;
```

```
RSQRRRGRRGEPGTALLVPLAIGLGLALACLGLLLAVVSLGSRASLSAQEPAQEBL 95
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
RRSQRRRGRRGEPGTALLVPLAIGLGLALACLGLLLAVVSLGSRASLSAQEPAQEBL 60

EDQDPSELNPQTEESQDPAPPLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQD 155
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
EDQDPSELNPQTEESQDPAPPLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQD 120
```

```
CC   -!- FUNCTION: Cytokine that binds to TNFRSF5. Mediates B-cell
CC       proliferation in the absence of co-stimulus as well as IgE
CC       production in the presence of IL-4. Involved in immunoglobulir
CC       class switching (By similarity).
CC   -!- SUBUNIT: Homotrimer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein. Also exists as
CC       extracellular soluble form (By similarity).
CC   -!- PTM: The soluble form derives from the membrane form by
CC       proteolytic processing (By similarity).
CC   -!- SIMILARITY: Belongs to the tumor necrosis factor family.
CC   --------------------------------------------------------------------------
```

```
DR   EMBL; AJ243435; CAB95748.2; -.
DR   HSSP; P29965; 1ALY.
DR   GO; GO:0016021; C:integral to membrane; ISS.
DR   GO; GO:0005174; F:CD40 receptor binding; ISS.
DR   GO; GO:0042100; P:B-cell proliferation; ISS.
DR   GO; GO:0006954; P:inflammatory response; ISS.
DR   GO; GO:0007159; P:leukocyte cell adhesion; ISS.
DR   GO; GO:0030168; P:platelet activation; ISS.
DR   InterPro; IPR003263; TNF_5.
DR   InterPro; IPR006052; TNF_family.
DR   InterPro; IPR008983; TNF_like.
DR   InterPro; IPR003636; TNF_subf.
DR   Pfam; PF00229; TNF; 1.
DR   PRINTS; PR01702; CD40LIGAND.
DR   ProDom; PD008600; TNF_5; 1.
DR   ProDom; PD002012; TNF_subf; 1.
DR   SMART; SM00207; TNF; 1.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 1.
KW   Cytokine; Transmembrane; Glycoprotein; Signal-anchor.
FT   CHAIN         1    272       TUMOR NECROSIS FACTOR LIGAND SUPERFAI
FT                                MEMBER 5, MEMBRANE FORM.
FT   CHAIN       111    272       TUMOR NECROSIS FACTOR LIGAND SUPERFAI
FT                                MEMBER 5, SOLUBLE FORM (BY SIMILARITY
FT   DOMAIN        1     23       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     24     44       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROT
FT                                (POTENTIAL).
FT   DOMAIN       45    272       EXTRACELLULAR (POTENTIAL).
FT   SITE        110    111       CLEAVAGE (BY SIMILARITY).
FT   DISULFID    190    229       POTENTIAL.
FT   CARBOHYD    124    124       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    146    146       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    251    251       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   272 AA;  30832 MW;  8CD0338A924E044B CRC64;
```

```
Query Match        7.6%;   Score 109.5;  DB 1;  Length 272;
Best Local Similarity 26.4%;   Pred. No. 0.14;  Indels 33;  G
Matches 42;  Conservative 26;  Mismatches 58;           Indels 33;  G

Qy   139  RRAIAAHYEVHPRPGQDGAQAGVDGTVSGW-EEARINSSSPLRYNRQIGEFIVTRA
              :||:|| || |:|   :: :||:
Db   133  RQPIATHLA------GVKSNTTVRVLKWMTTSYAPTSSLISYHE--GKLKVEKA

Qy   198  LYCQVHFDEGKA-----------VYLKLDLLVDGVLALRCLEEFSATAASSLGPQLRLC
              :||:|:| |          :::| ||    |:::
Db   183  IYSQVSFCTKAAASAPFTLYIYLPMEEDRLL-MKGLDTHSTSTA-------LC

Qy   248  --SGLLALRPGSSLRIRTLPWAHLKAAPFLTYFGLFQV 283
             ||:|::||||||||::||::
Db   234  IREGGVFELRQGDMVFVNVTDSTAVNVPGNTYFGMFKL 272

RESULT 4
TNF5_CANFA
```

```
--------------------------------------------------------------------------
;  AAC53517.1; -.
;  BAB32249.1; -.
59; Tnfsf12.
06052; TNF_family.
08983; TNF_like.
TNF; 1.
51; TNF_1; FALSE_NEG.
49; TNF_2; 1.
ogenesis; Apoptosis; Transmembrane; Glycoprotein;

1        1               TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
1        225             MEMBER 12, MEMBRANE FORM.
0        225             TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
                         MEMBER 12, SECRETED FORM (BY SIMILARITY).
1        21              SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
                         (POTENTIAL).
2        225             EXTRACELLULAR (POTENTIAL).
9        70              CLEAVAGE (BY SIMILARITY).
7        186             POTENTIAL.
5        115             N-LINKED (GLCNAC...) (POTENTIAL).
AA;  24781 MW;  90C412CC04B0659B CRC64;

rity        70.6%;   Score 1020;  DB 1;  Length 225;
            88.8%;   Pred. 8.5e-73;
onservative  9;  Mismatches 16;  Indels 0;  Gaps 0;

ALACLGLLLAVVSLGSRASLSAQEPAQELVAEEDQDPSELNPQTEESQDPAPFL  120
|||||||||:|||||: :||||||||||: ::||||||||||||:||||
ALACLGLLLVVVSLGSWATLSAQEPSQEELTAEDRREPPELNPQTEESQDVVPFL  61

PRRSAPKGRKTRARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLR 180
||||||||||||||||||||||||||||||||||||||||: :|||||||||||
PRRSAPKGRKARPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSSSPLR 121

GEFIVTRAGLYYLYCQVHFDEGKAVYLKLDLLVDGVLALRCLEEFSATAASSLGP 240
|||::||||||||||||||||||||||||||||||||||||||||||||||||
GEFVIRAGLYYLYCQVHFDEGKAVYLKLDLLVNGVLALRCLEEFSATAASSPGP 181

QVSGLLALRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 284
||||||||||||||||||||||||||||||||||||||
QVSGLLPLRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 225

STANDARD;      PRT;   272 AA.

Rel. 41, Created)
Rel. 43, Last sequence update)
Rel. 43, Last annotation update)
s factor ligand superfamily member 5 (CD40 ligand) (CD40-
ein).
DLG OR CD40L.
(Chicken).
:azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
:ves; Neognathae; Galliformes; Phasianidae; Phasianinae;
31;

N.A.
leghorn; TISSUE=Spleen;
, Young J.R., Burnside J.;
, Burnside J.;
putative chicken CD40 ligand.";
R-2003) to the EMBL/GenBank/DDBJ databases.
```

```
Db   103 EMKKEENIAMQKGDQDPRIAAHVISEASSNPASVL------RWAPKGYYTISSNI
             |:||::||::|:|     ||||| :     ::|    |:|      |||||
Qy   146 YEVHPRPCQDGAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCC
                                                        ::||
Db   155 ------ENGKQ------                                  LAVKRQGLYYVYAC
                                                            :|||:|
Qy   206 EGKAVYLKLDLLVDGVLALRCLEEFSAT-----AASSLGPQLRLCQVS----GLI
                              ::|      :     ||||| ||:::      :|
Db   178 SNRAASSQAPF-----VASLCLHSPGGTERVLLRAASSRGSSKPCGQQSIHLGGVF

Qy   256 GSSLRIRTLPWAHLKAAPFLTYFGLFQV 283
             |::|::    ||||| :::::
Db   233 GASVFVNVTDFSQVSHGTGFTSFGLLKL 260
```

RESULT 5

TNF9_HUMAN

ID   TNF9_HUMAN   STANDARD;   PRT;   254 AA.
AC   P41273;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Tumor necrosis factor ligand superfamily member 9 (4-1BB ligand)
DE   1BBL).
GN   TNFSF9.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94374434; PubMed=8088337;
RA   Alderson M.R., Smith C.A., Tough T.W., Davis-Smith T., Armitage F
RA   Falk B., Roux E., Baker E., Sutherland G.R., Din W.S., Goodwin R.
RT   "Molecular and biological characterization of human 4-1BB and its
RT   ligand.";
RL   Eur. J. Immunol. 24:2219-2227(1994).
CC   -!- FUNCTION: Cytokine that binds to TNFRSF9. Induces the
CC   proliferation of activated peripheral blood T cells. May have
CC   role in activation-induced cell death (AICD). May play a role
CC   cognate interactions between T cells and B cells/macrophages.
CC   -!- SUBUNIT: Homotrimer (Potential).
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, PLACENTA, LUNG, SKELI
CC   MUSCLE AND KIDNEY.
CC   -!- SIMILARITY: Belongs to the tumor necrosis factor family.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a col
CC   between the Swiss Institute of Bioinformatics and the EMBL out
CC   the European Bioinformatics Institute. There are no restrictio
CC   use by non-profit institutions as long as its content is i
CC   modified and this statement is not removed. Usage by and for (
CC   entities requires a license agreement (See http://www.isb-sib.ch,
CC   or send an email to license@isb-sib.ch).
DR   EMBL; U03398; AAA53134.1; -.
DR   PIR; I38427; I38427.
DR   Genew; HGNC:11939; TNFSF9.
DR   MIM; 606182; -.
DR   GO; GO:0005102; F:receptor binding; TAS.
DR   GO; GO:0006915; P:apoptosis; TAS.
DR   GO; GO:0008283; P:cell proliferation; TAS.
DR   GO; GO:0007267; P:cell-cell signaling; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR006052; TNF family.
DR   InterPro; IPR008983; TNF_like.
DR   Pfam; PF00229; TNF; 1.
DR   SMART; SM00207; TNF; 1.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 1.
KW   Cytokine; Transmembrane; Glycoprotein; Signal-anchor; Polymorphi:
FT   DOMAIN   1   28   CYTOPLASMIC (POTENTIAL).
```

```
Db

Qy

Db

Qy

Db

Qy

Db
```

STANDARD;   PRT;   260 AA.

Rel. 40, Created)
Rel. 40, Last sequence update)
Rel. 41, Last annotation update)
s factor ligand superfamily member 5 (CD40 ligand).
OLG OR CD40L.
ris (Dog).
tazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
heria; Carnivora; Fissipedia; Canidae; Canis.
15;

N.A.
illett B.J.;
perties of canine CD40L.";
G-1998) to the EMBL/GenBank/DDBJ databases.
Cytokine that binds to TNFRSF5. Mediates B-cell
tion in the absence of co-stimulus as well as IgE
n in the presence of Il-4. Involved in immunoglobulin
tching (By similarity).
Homotrimer (By similarity).
AR LOCATION: Type II membrane protein. Also exists as an
ular soluble form (By similarity).
soluble form derives from the membrane form by
ic processing (By similarity).
Y: Belongs to the tumor necrosis factor family.

OT entry is copyright. It is produced through a collaboration
Swiss Institute of Bioinformatics and the EMBL outstation -
Bioinformatics Institute. There are no restrictions on its
profit institutions as long as its content is in no way
this statement is not removed. Usage by and for commercial
ires a license agreement (See http://www.isb-sib.ch/announce/
ail to license@isb-sib.ch).

1; AAD04375.1; -.
IALY.
1; C:integral to membrane; ISS.
4; F:CD40 receptor binding; ISS.
10; P:B-cell proliferation; ISS.
4; P:inflammatory response; ISS.
9; P:leukocyte cell adhesion; ISS.
8; P:platelet activation; ISS.
003263; TNF_5.
006052; TNF_family.
008983; TNF_like.
003636; TNF_subf.
; TNF; 1.
02; CD40LIGAND.
600; TNF_5; 1.
012; TNF_subf; 1.
7; TNF_1.
251; TNF_1; 1.
049; TNF_2; 1.
unsmembrane; Glycoprotein; Signal-anchor.
1   260   TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
                    MEMBER 5, MEMBRANE FORM.
12   260   TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
                    MEMBER 5, SOLUBLE FORM (BY SIMILARITY).
1   22   CYTOPLASMIC (POTENTIAL).
23   46   SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
                    (POTENTIAL).
47   260   EXTRACELLULAR (POTENTIAL).
111   112   CLEAVAGE (BY SIMILARITY).
177   217   POTENTIAL.
239   239   N-LINKED (GLCNAC...) (POTENTIAL).
50 AA; 28688 MW; 604F69A19E98EB70 CRC64;
larity   7.5%; Score 108.5; DB 1; Length 260;
             Pred. No. 0.16;
conservative 25.5%; Mismatches 69; Indels 63; Gaps   8;
onservative 23;   Mismatches 23;

EELVAEE--DQDPS-ELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAH 145
```

9    .    49    SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
                       (POTENTIAL).
0         254    EXTRACELLULAR (POTENTIAL).
5          41    POLY-LEU.
7          17    P -> A (in dbSNP:442511).
                 /FTId=VAR_011928.

AA; 26624 MW; 827551F34563E508 CRC64;

rity    7.4%; Score 106.5; DB 1; Length 254;
26.5%; Pred. No. 0.22;
nservative 31; Mismatches 93; Indels 81; Gaps 13;

AARRSQRRGRRGEPGTALLVPLAGLGLALACLGLLLAVVSL-GSRASL-SAQE 89
     |||: ||: ||    :|:|||||||||||    :: :|||  :|:|||:
----RARACRVLP-WALVAGLLLLLLAAACAVFLACPWAVSGARASPGSAAS 67

LVAEEQDPSELNPQTEBSQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVH 149
    :|:|                         :  :|:     :|
------GPELSP-----DDPAGLLDL----------RQGMFAQLVAQNVLL- 102

DGAQAGVDGTVSGWEE--ARINSSSPLRYRNRQIGEFIVTRAGLYYLYCQVHFDE 206
   :||:|:  |:|    :|  |: |:|||||  ::|  |  |:|:|::|||::|
-----IDGPLSWYSDPGLAGVSLTGGLSYKEDTKElVVAKAGVYYVFFQ----- 146

LKLDLLVDG------VLALRCLEEFSATAASSLGPQLRLCQVSG----------- 249
|:| :|          |||   |:|         |:|||   :|
LELRRVVAGEGGSGVSLALHLQPLRSAAGAAALALTVDLPPASSEARNSAFGFQG 201

RPGSSLRIRTLPWAHLKAAPFL----TYFGLFQV 283
  |:|       |:| ||            ||:||
SAGQRLGVHLHTEARARHAWQLTQGATVLGLFRV 240

STANDARD;    PRT;    952 AA.
Q96K01; Q9BR73; Q9H7L0; Q9NW41; Q9NWA9; Q9NYK9;

el. 42, Created)
el. 42, Last sequence update)
el. 43, Last annotation update)
ylase 7A (HD7a).
7.
azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
eria; Primates; Catarrhini; Hominidae; Homo.
6;

N.A. (ISOFORM 1).
1 carcinoma;
W., Verdin E., Walsh M.J.;
II HDAC is associated with the transcriptional
pressor CCAAT displacement protein.";
-2000) to the EMBL/GenBank/DDBJ databases.

N.A. (ISOFORM 2), SEQUENCE OF 220-952 FROM N.A. (ISOFORM
CE OF 651-952 FROM N.A.
Placenta, Spleen, and Teratocarcinoma;
T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
Nagai K., Sugano S., Shiratori A., Sudo H.,
Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
Chiba J., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
imura K., Murakami K., Ishii S., Kawai Y., Saito K.,
akamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
wayanagi T.;
NA sequencing project.";
-2003) to the EMBL/GenBank/DDBJ databases.

N.A. (ISOFORM 3).
.; histone deacetylase 7A (HDAC7A), transcript variant 3.";
-2003) to the EMBL/GenBank/DDBJ databases.

RP   SEQUENCE FROM N.A.
RA   Worley K.C.;
RL   Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE OF 75-952 FROM N.A. (ISOFORM 1).
RC   TISSUE=Uterus;
RA   Koehrer K., Beyer A., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE OF 241-952 FROM N.A. (ISOFORM 1).
RC   TISSUE=B-cell, and Colon;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Rodrigues S., Sanchez
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length h
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [7]
RP   SUBCELLULAR LOCATION, AND INTERACTION WITH EDNRA.
RX   MEDLINE=21264398; PubMed=11262386;
RA   Lee H.-J., Chun M., Kandror K.V.;
RT   "Tip60 and HDAC7 interact with the endothelin receptor a and may b
RT   involved in downstream signaling.";
RL   J. Biol. Chem. 276:16597-16600(2001).
RN   [8]
RP   INTERACTION WITH HDAC3.
RX   MEDLINE=21443773; PubMed=11466315;
RA   Fischle W., Dequiedt F., Fillion M., Hendzel M.J., Voelter W.,
RA   Verdin E.;
RT   "Human HDAC7 histone deacetylase activity is associated with HDAC3
RT   vivo.";
RL   J. Biol. Chem. 276:35826-35835(2001).
RN   [9]
RP   FUNCTION.
RX   MEDLINE=22224741; PubMed=12239305;
RA   Bryant H., Farrell P.J.;
RT   "Signal transduction and transcription factor modification during
RT   reactivation of Epstein-Barr virus from latency.";
RL   J. Virol. 76:10299-10298(2002).
RN   [10]
RP   INTERACTION WITH HTATIP.
RX   MEDLINE=22538515; PubMed=12551922;
RA   Xiao H., Chung J., Kao H.-Y., Yang Y.-C.;
RT   "Tip60 is a co-repressor for STAT3.";
RL   J. Biol. Chem. 278:11197-11204(2003).

-!- FUNCTION: Responsible for the deacetylation of lysine residues
CC       the N-terminal part of the core histones (H2A, H2B, H3 and H4)
CC       Histone deacetylation gives a tag for epigenetic repression an
CC       plays an important role in transcriptional regulation, cell cy
CC       progression and developmental events. Histone deacetylases act
CC       the formation of large multiprotein complexes. Involved in mus
CC       maturation by repressing transcription of myocyte enhancer fac
CC       such as MEF2A, MEF2B and MEF2C. During muscle differentiation,
CC       shuttles into the cytoplasm, allowing the expression of myocyt
CC       enhancer factors (By similarity). May be involved in Epstein-B
CC       virus (EBV) latence, possibly by repressing the viral BZLF1 ge
-!- SUBUNIT: Interacts with HDAC1, HDAC2, HDAC3, HDAC4, HDAC5, NC
CC       NCOR2, SIN3A, SIN3B, RBBP4, RBBP7, MTA1L1, SAP30 and MBD3.

```
FT   DOMAIN      218   546   TRANSCRIPTION REPRESSION 2 (BY
FT                           SIMILARITY).
FT   DOMAIN      518   865   HISTONE DEACETYLASE.
FT   DOMAIN      918   952   NUCLEAR EXPORT (BY SIMILARITY).
FT   DOMAIN        1    98   INTERACTION WITH MEF2C (BY SIMILARIT
FT   DOMAIN       49   149   INTERACTION WITH MEF2A (BY SIMILARIT
FT   DOMAIN      877   952   INTERACTION WITH SIN3A (BY SIMILARIT
FT   DOMAIN      197   203   POLY-SER.
FT   DOMAIN      368   373   POLY-PRO.
FT   ACT_SITE    670   670   BY SIMILARITY.
FT   VARSPLIC      1   472   Missing (in isoform 2).
FT                           /FTId=VSP_007429.
FT   VARSPLIC    227   263   Missing (In isoform 3).
FT                           /FTId=VSP_008772.
FT   VARSPLIC    227   256   Missing (In isoform 4).
FT                           /FTId=VSP_007430.
```

```
Query Match     6.9%;  Score 99;  DB 1;  Length 952;
Best Local Similarity   24.2%;  Pred. No. 3.9;
Matches  46;  Conservative  21;  Mismatches  67;  Indels  56;  G

Qy    11 RRLPLPRSLGSRDGAVR----------------QAQPPAPMAAR------RSQRR
         |::|.|          :|. |||
Db   401 RQIPSAEDLETDGGPGQVVDDGLEHRELGHGQPEARGPAPLQQHPQVLLWEQQRI

Qy    48 -RGEPGTALLVPIALGLGLALACIGLLLAVVSLGSRASLSAQEPPAQEELVAEEDQI
          :|
Db   461 PRGSTGDTVLLPLAQGGHRPLS-----RAQSSPAAPASLSAPEPASQARVLSSSET

Qy   107 NPQT-----------EESQDP----------APFLNRLVRPRRSAPKGR#
         : |:                    :|.    |||
Db   516 LPFTTGLIYDSVMLKHQCSGDNSRHPEHAGRIQSIWSRLQERGLRSQCECLRGR#

Qy   140 RAIAAHYEVH 149
         |.||
Db   576 ELQSVHSERH 585
```

```
RESULT 7
CG22_ANTMA
ID   CG22_ANTMA   STANDARD;   PRT;   441 AA.
AC   P34801;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   G2/mitotic-specific cyclin 2.
OS   Antirrhinum majus (Garden snapdragon).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; aste
OC   lamiids; Lamiales; Antirrhineae; Antirrhinum.
OX   NCBI_TaxID=4151;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94148008; PubMed=8313906;
RA   Fobert P.R., Coen E.S., Murphy G.J.P., Doonan J.H.;
RT   "Patterns of cell division revealed by transcriptional regulatio
RT   genes during the cell cycle in plants.";
RL   EMBO J. 13:616-624(1994).
CC   -!- FUNCTION: Essential for the control of the cell cycle at the
CC       (mitosis) transition. G2/M cyclins accumulate steadily durin
CC       and are abruptly destroyed at mitosis.
CC   -!- SUBUNIT: Interacts with the CDC2 and CDK2 protein kinases to
CC       a serine/threonine kinase holoenzyme complex. The cyclin sub
CC       imparts substrate specificity to the complex.
CC   -!- DEVELOPMENTAL STAGE: Accumulates steadily during G2 and is
CC       abruptly destroyed at mitosis.
CC   -!- SIMILARITY: Belongs to the cyclin family. Cyclin AB subfamil
CC
CC   This SWISS-PROT entry is copyright. It is produced through a col
CC   between the Swiss Institute of Bioinformatics and the EMBL out
CC   the European Bioinformatics Institute. There are no restrictio
CC   use by non-profit institutions as long as its content is li
CC   modified and this statement is not removed. Usage by and for
```

```
with the 14-3-3 protein YWHAE, MEF2A, MEF2B and MEF2C
arity). Interacts with HTATIP and EDNRA.
AR LOCATION: Nuclear and cytoplasmic. In the nucleus, it
s with distinct subnuclear dot-like structures. Shuttles
ne nucleus and the cytoplasm. Treatment with EDN1 results
ing from the nucleus to the perinuclear region. The
cytoplasm depends on the interaction with the 14-3-3
WHAE and may be due to its phosphorylation.
VE PRODUCTS:
ernative splicing; Named isoforms=4;

8WUI4-1; Sequence=Displayed;

8WUI4-2; Sequence=VSP_007429, VSP_007431;
experimental confirmation available;

8WUI4-4; Sequence=VSP_008772;

8WUI4-3; Sequence=VSP_007430;
experimental confirmation available;

he nuclear export sequence mediates the shuttling between
us and the cytoplasm (By similarity).
be phosphorylated by CaMK1 (By similarity).
EOUS: Its activity is inhibited by Trichostatin A (TSA),
istone deacetylase inhibitor (By similarity).
Y: Belongs to the histone deacetylase family. Subfamily

Ref.1 sequence differs from that shown due to a
t in position 877.
Ref.2 (BAC56929) sequence differs from that shown due to
etention.

OT entry is copyright. It is produced through a collaboration
Swiss Institute of Bioinformatics and the EMBL outstation -
Bioinformatics Institute. There are no restrictions on its
profit institutions as long as its content is in no way
this statement is not removed. Usage by and for commercial
ires a license agreement (See http://www.isb-sib.ch/announce/
ail to license@isb-sib.ch).

3;  AAF63491.1; ALT_FRAME.
2;  BAA91474.1; ALT_INIT.
0;  BAA91545.1; ALT_INIT.
9;  BAB15759.1; ALT_INIT.
1;  BAB55363.1; ALT_INIT.
8;  BAC56929.1; ALT_SEQ.
7;  AAP84704.1; -.
6;  -; NOT ANNOTATED_CDS.
5;  CAB55935.1; -.
3;  AAH06453.1; ALT_INIT.
5;  AAH20505.1; ALT_INIT.
4067; HDAC7A.

.7;  C:cytoplasm; TAS.
.8;  C:histone deacetylase complex; TAS.
.4;  C:nucleus; TAS.
07;  F:histone deacetylase activity; TAS.
6;   F:specific transcriptional repressor activity; TAS.
4;   F:transcription factor binding; TAS.
3;   P:B-cell differentiation; TAS.
8;   P:chromatin modification; TAS.
4;   P:inflammatory response; TAS.
3;   P:negative regulation of myogenesis; TAS.
9;   P:neurogenesis; TAS.
4;   P:regulation of cell cycle; TAS.
00286; His_deacetylase.
0;   Hist_deacetyl; 1.
270; HDASUPER.

clear protein; Chromatin regulator;
regulation; Repressor; Repeat; Phosphorylation;
splicing;
1     268    TRANSCRIPTION REPRESSION 1 (BY
             SIMILARITY).
```

CAA53729.1; -.
1VIN.
41710.
06670; Cyclin.
04367; Cyclin_Cterm.
06671; Cyclin_N.
cyclin; 1.
cyclin_C; 1.
; CYCLIN; 2.
92; CYCLINS; 1.
ycle: Cell division: Mitosis.
AA; 49205 MW; E6E4C037C98880A7 CRC64;

rity   6.7%;   Score 97;   DB 1;   Length 441;
         24.0%;   Pred. No. 2.3;   Indels   104;   Gaps   15;
nservative   44;   Mismatches   74;

AQPPAPMAARRSQRR---------RGRGEPGTALLVPLAIGLGLALACLGLL 73
                        || || |: |: | |:
KS----MAVEKKNRALGDIGNVVTVRGVEGKALPQVSRPITRGF-----CAQLI 69

--------LAVVSLGSRASLS------AQEPAQELVAEEDQDPSELNPQTEBSQ 114
         ||| |:| | |: ||:||:|:|||||:| |:|: |||:
AAAAENNKNSLAVNAKGADGALPIKRAVARVPVQKKTVKSKPQBIIEISPDTEKKK 129

LNRLVRPRRS------APKGRKTRARRAIAAHYEVHPRGQ-----DGAQAGVDGTV 165
  ||| || | |||| |||:   :: |
LEKBITGEKSLKKKAPTLTSTLTARSKAASV-VRTKPKBQIVDIDAADVNNDLAV 186

---ARINSSSPLRY--NRQIGEFIVTRAGLYYLYCQVHFD--EGKAVYL 212
    | | |||| ||:|||||||||||| |:| | ||
DMYKFYKSAENDSRPHDYMDSQPEINEKM--RAILIDWLVQVHYKFELSPETLYL 244

VDGVLALRC------LEEFSATAASSLGPQLRLCQVSGLLALRPGS 257
 || | ::| || | || :| |:| |: :|:| ||
VDRYLASKITSRRELQLLGMSSMLIASKYEEIWAPEVNDLVCISDGS 295

STANDARD;   PRT;   201 AA.

Rel. 40, (Created)
Rel. 40, Last sequence update)
Rel. 42, Last annotation update)
pha precursor (LT-alpha) (TNF-beta) (Tumor necrosis
superfamily member 1).
OR TNFB.
ii (Tammar wallaby).
azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
theria; Diprotodontia; Macropodidae; Macropus.
5;
N.A.
48; PubMed=10826697;
Deane E.M.;
of lymphotoxin alpha (LT-alpha) from a marsupial,
iii.;
09-403(2000).
Cytokine that in its homotrimeric form binds to
TNFR1, TNFRSF1B/TNFBR and TNFRSF14/HVEM. In its
meric form with LTB binds to TNFRSF3/LTBR. Lymphotoxin is
by lymphocytes and cytotoxic for a wide range of tumor
vitro and in vivo.
Homotrimer, and heterotrimer of either two LTB and one
its or (less prevalent) two LTA and one LTB subunits (By
/).
R LOCATION: Secreted (homotrimer) and membrane-
d (heterotrimers) (By similarity).

DR   EMBL; AF119336; AAD41773.1; -.
DR   HSSP; P01374; 1TNR.
DR   InterPro; IPR006053; TNF_abc.
DR   InterPro; IPR006052; TNF_family.
DR   InterPro; IPR008983; TNF_like.
DR   InterPro; IPR003636; TNF_subf.
DR   Pfam; PF00229; TNF; 1.
DR   PRINTS; PR01234; TNECROSISFCT.
DR   ProDom; PD002012; TNF_subf; 1.
DR   SMART; SM00207; TNF; 1.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 1.
KW   Cytokine; Glycoprotein; Signal.
FT   SIGNAL   1   27   BY SIMILARITY.
FT   CHAIN   28   201   LYMPHOTOXIN-ALPHA.
FT   CARBOHYD   93   93   N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   201 AA;   21536 MW; BC4C371CB5091627 CRC64;

Query Match   6.6%;   Score 95.5;   DB 1;   Length 201;
Best Local Similarity 23.2%;   Pred. No. 1.2;
Matches   44;   Conservative   27;   Mismatches   84;   Indels   35;   G:

Qy   107 NPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAI--AAHYEVHPRPGQDGAQAG:
          || |||||||||||||||||||||| |: ||| |||:
Db   30   NPDNSHSSSPAP-------PQTAQHLSQKSLKRETLKPAAHL-----VGDPSVQDS:

Qy   165 VSGWEEARINSSSP-LRYNRQI--GEFIVTRAGLYYLYCQVHFDEGKA--------V
          ||: | ||||||||||||| |||| |
Db   76   ---W--RANTDHAFLRHGFSLSNNSLLVPTSGLYFVYSQVVFSGASCSEITPTLL:

Qy   215 DLLVDG---VLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWA:
          ::|: |:| || | | ||| :| |:| |: :|:|
Db   130 EVLLFSSKYQVHVPLLSAQKSVCGSTQGPWMRSVYQGAVFLLTQGDRLSTYTDGVS:

Qy   272 APFLTYFGLF 281
          :| ||||:
Db   190 SPSSVFFGAF 199

RESULT 9
TNF6_MACMU
ID   TNF6_MACMU   STANDARD;   PRT;   280 AA.
AC   Q9MYL6; Q9BDM5;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Tumor necrosis factor ligand superfamily member 6 (FAS antigen li:
DE   (CD95L protein).
GN   TNFSF6 OR FASL OR CD95L.
OS   Macaca mulatta (Rhesus macaque).
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), an
OS   Macaca nemestrina (Pig-tailed macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544, 9541, 9545;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   SPECIES=M.mulatta; TISSUE=Lymphocytes;
RX   MEDLINE=21383618; PubMed=11491535;
RA   Villinger F., Bostik P., Mayne A.E., King C.L., Genain C.P.,
RA   Weiss W.R., Ansari A.A.;
RT   "Cloning, sequencing, and homology analysis of nonhuman primate
RT   Fas/Fas-ligand and co-stimulatory molecules.";
RL   Immunogenetics 53:315-328(2001).

Db   79 GLCLLVMFFMVLVALVGLG--LGMFQLFHLQKEL------AELRESTSQKHTAS

Qy   122 RLVRPRRSAPKGRKTRARATAAHYEVHPRPGQDGAQAGVDGTVSGWEEA-RINSS
              :: : |:|||:|:|:: |::: ::|

Db   129 QIGHP--SPPPEKKEQRK--VAHLTGKPNSRSMPLE------WEDTYGIVLL

Qy   181 YNRQIGEFIVTRAGLIYLYCQVHFDEGKA------VYLKLD-----LLVDGV
              : : ::|||:::|:|:|:| :|:   ::|:

Db   175 YKK--GGLVINETGLIFVYSKVYF-RGQSCTNLPLSHKVYMRNSKYPQDLVMMEGK
              :||| ::| | |

Qy   226 CLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFLYYFGLF
              ::|| :|| :| : |: : |

Db   232 CTTGQMWAHSSYIGAVFNLTSADHLY------VNVSELSLVNFEESQ--TFFGLY

RESULT 10

RHO_MICLU
ID   RHO_MICLU   STANDARD;   PRT;   690 AA.
AC   P52154;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Transcription termination factor rho.
GN   RHO.
OS   Micrococcus luteus (Micrococcus lysodeikticus).
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Micrococcineae; Micrococcaceae; Micrococcus.
OX   NCBI_TaxID=1270;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 1-5 AND 289-297.
RC   STRAIN=EM;
RX   MEDLINE=96132802; PubMed=8857681;
RA   Nowatzke W.L., Richardson J.P.;
RT   "Characterization of an unusual Rho factor from the high G + C gr
RT   positive bacterium Micrococcus luteus.";
RL   J. Biol. Chem. 271:742-747(1996).
RN   [2]
RP   SEQUENCE OF 205-690 FROM N.A.
RC   STRAIN=EM;
RX   MEDLINE=94327472; PubMed=8051015;
RA   Opperman T., Richardson J.P.;
RT   "Phylogenetic analysis of sequences from diverse bacteria with
RT   homology to the Escherichia coli rho gene.";
RL   J. Bacteriol. 176:5033-5043(1994).
RN   [3]
RP   REVISION TO 500.
RC   STRAIN=EM;
RA   Nowatzke W.L.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: FACILITATES TRANSCRIPTION TERMINATION BY A MECHANIS
CC   THAT INVOLVES RHO BINDING TO THE NASCENT RNA, ACTIVATION OF I
CC   RNA-DEPENDENT ATPASE ACTIVITY, AND RELEASE OF THE MRNA FROM T
CC   DNA TEMPLATE. RNA-DEPENDENT NTPASE WHICH UTILIZES ALL FOUR
CC   RIBONUCLEOSIDE TRIPHOSPHATES AS WELL AS DATP AS SUBSTRATES, B
CC   HAS A SIGNIFICANT LOWER ACTIVITY WITH CTP.
CC   -!- SUBUNIT: Homohexamer (By similarity).
CC   -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC
CC   This SWISS-PROT entry is copyright. It is produced through a col
CC   between the Swiss Institute of Bioinformatics and the EMBL out
CC   the European Bioinformatics Institute. There are no restrictio
CC   use by non-profit institutions as long as its content is mo
CC   modified and this statement is not removed. Usage by and for c
CC   entities requires a license agreement (See http://www.isb-sib.ch/
CC   or send an email to license@isb-sib.ch).
DR   EMBL; L27277; AAB18671.1; -.
DR   HSSP; P03002; 1A63.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR000194; ATPase_a/bcentre.
DR   InterPro; IPR002059; Cold_shock.
DR   InterPro; IPR008994; Nucleic_acid_OB.
DR   InterPro; IPR004665; Term_rho.

N.A.
atta, M.fascicularis, and M.nemestrina;
ue T., Yoshino K.;
V-1999) to the EMBL/GenBank/DDBJ databases.
Cytokine that binds to TNFRSF6/FAS, a receptor that
s the apoptotic signal into cells. May be involved in
T cell mediated apoptosis and in T cell development.
AS-mediated apoptosis may have a role in the induction of mature
l tolerance, in the antigen-stimulated suicide of mature
or both. Binding to the decoy receptor TNFRSF6B/DcR3
its effects (By similarity).
Homotrimer (Potential).
AR LOCATION: Type II membrane protein and secreted (By
y).
soluble form derives from the membrane form by
ic processing (By similarity).
Y: Belongs to the tumor necrosis factor family.

OT entry is copyright. It is produced through a collaboration
Swiss Institute of Bioinformatics and the EMBL outstation -
Bioinformatics Institute. There are no restrictions on its
profit institutions as long as its content is in no way
this statement is not removed. Usage by and for commercial
ires a license agreement (See http://www.isb-sib.ch/announce/
ail to license@isb-sib.ch).

6; AAK37539.1; -.
8; BAA90294.1; -.
9; BAA90295.1; -.
0; BAA90296.1; -.
4TSV.
008064; Fas ligand.
006053; TNF_abC.
006052; TNF_family.
008983; TNF_like.
003636; TNF_subf.
; TNF; 1.
81; FASLIGAND.
34; TNECROSISFCT.
012; TNF_subf; 1.
7; TNF; 1.
251; TNF_1; 1.
049; TNF_2; 1.
ptosis; Transmembrane; Glycoprotein; Signal-anchor.
1   280   TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
                MEMBER 6, MEMBRANE FORM.
29   280   TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
                MEMBER 6, SOLUBLE FORM (BY SIMILARITY).
1   80   CYTOPLASMIC (POTENTIAL).
81   101   SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
                (POTENTIAL).
02   280   EXTRACELLULAR (POTENTIAL).
4   69   PRO-RICH.
45   64   POLY-PRO.
28   129   CLEAVAGE (BY SIMILARITY).
01   232   POTENTIAL.
83   183   N-LINKED (GLCNAC...) (POTENTIAL).
49   249   N-LINKED (GLCNAC...) (POTENTIAL).
59   259   N-LINKED (GLCNAC...) (POTENTIAL).
60   60   S -> P (IN REF. 1).
30 AA; 31367 MW; F0B284D61A132EB4 CR C64;

larity: 6.5%; Score 94; DB 1; Length 280;
onservative 20.8%; Pred. No. 2.4; Mismatches 117; Indels 74; Gaps 14;
             45; Mismatches 117; Indels 74; Gaps 14;

RSLGSRDGGAVRQAQPPAPMAAARRSQRRRGRRGEPGTALLVPLAL------------- 61
  |:|:|| ||| || :
TSVPRRPGQRRPPPPPPPLP-----------PPPPSPLPPLPLPPLKKRGNHST 78
  :||| || :| || |

ALACLGLLLAVVSLGRSRSLSAQEPAQEELVAEEDQDPSELNPQTEESQDPAPFLN 121
  |:|:||| || || :| ||

ATP-synt_ab; 1.
; AAA; 1.
; CSP; 1.
00767; rho; 1.
ermination; Helicase; ATP-binding; RNA-binding.

```
0         0    RNA-BINDING (RNP2)  (BY SIMILARITY).
0        35    RNA-BINDING (RNP1)  (BY SIMILARITY).
3       321    ATP (POTENTIAL).
1       448    G -> P (IN REF. 2; AA SEQUENCE).
1       291
AA;   75030 MW;  F77C4C75E1B8998 CRC64;
```

rity   6.5%; Score 93.5; DB 1; Length 690;
nservative 25.9%; Pred. No. 7.3; Mismatches 63; Indels 41; Gaps 7;
nservative 19; Mismatches

EISARRLPLPRSLGSRDGGAVRQAQPPAPMAARRSQRRRGRRGEPGTALLVPLAL 61
:   ::::::|||
DAAERAAQAPAAPAAETAPAAASSEDAAP-AAERPARRRSRRADADTS--APAAA 111

LACLGLLLAVVSLGSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQDPAPFLN 121
: : ::| || | ||| ||
QA---------EAREAQTEQAPRE----TASDQDRSGGSEARDEGED-------- 150

RRSAPKGRKTRARRAIAAHYEVHPRPGQDGAQAGVDGTVSG 167
: : :||||
QSE--RRSRGRR--------RAGDDDAQQGQDRRSDG 179

STANDARD;   PRT;   310 AA.

el. 34, Created)
el. 34, Last sequence update)
el. 42, Last annotation update)
rotein Rv0497/MT0517/MB0508.
17 OR MTCY20G9.23 OR MB0508.
tuberculosis, and
bovis.
nobacteria; Actinobacteridae; Actinomycetales;
eae; Mycobacteriaceae; Mycobacterium.

N.A.
rculosis; STRAIN=H37Rv;
87; PubMed=9634230;
sch R., Parkhill J., Garnier T., Churcher C., Harris D.,
iglmeier K., Gas S., Barry C.E. III, Tekaia F.,
sham D., Brown D., Chillingworth T., Connor R.,
lin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
gels K., Krogh A., McLean J., Moule S., Murphy L.,
orne J., Quail M.A., Rajandream M.A., Rogers J.,
ger K., Skelton S., Squares S., Squares R.,
Taylor K., Whitehead S., Barrell B.G.;
he biology of Mycobacterium tuberculosis from the
e sequence.";
-544(1998).

N.A.
rculosis; STRAIN=CDC 1551 / Oshkosh;
94; PubMed=12218036;
D., Alland D., Eisen J.A., Carpenter L., White O.,
eBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
terback T., Weidman J., Khouri H., Gill J., Mikula A.,
obs W.R. Jr., Venter J.C., Fraser C.M.;
comparison of Mycobacterium tuberculosis clinical and
ains.";
184:5479-5490(2002).

N.A.
s; STRAIN=AF2122/97;

RX   MEDLINE=22709107; PubMed=12788972;
RA   Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA   Pryor M., Duthoy S., Grondin S., Lacroix C., Monsempe C., Simon S.,
RA   Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA   Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT   "The complete genome sequence of Mycobacterium bovis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- SIMILARITY: TO M.LEPRAE ML2433.
CC
CC   This SWISS-PROT entry is copyright. It is produced through a colla
CC   between the Swiss Institute of Bioinformatics and the EMBL outs
CC   the European Bioinformatics Institute. There are no restriction
CC   use by non-profit institutions as long as its content is in
CC   modified and this statement is not removed. Usage by and for cc
CC   entities requires a license agreement (See http://www.isb-sib.ch/a
CC   or send an email to license@isb-sib.ch).
CC
DR   EMBL; Z77162; CAB00923.1; -.
DR   EMBL; AE006952; AAK44740.1; -.
DR   EMBL; BX248335; CAD93371.1; -.
DR   PIR; D70745; D70745.
DR   TIGR; MT0517; -.
DR   TubercuList; Rv0497; -.
KW   Hypothetical protein; Transmembrane; Complete proteome.
FT   TRANSMEM   231   251   POTENTIAL.
FT   TRANSMEM   257   277   POTENTIAL.
FT   TRANSMEM   286   306   POTENTIAL.
FT   DOMAIN   33   39   POLY-ARG.
FT   DOMAIN   197   202   POLY-ALA.
SQ   SEQUENCE   310 AA;   33092 MW; 4954027F694DF5C2 CRC64;

Query Match   6.4%;   Score 93; DB 1; Length 310;
Best Local Similarity 24.4%; Pred. No. 3.2;
Matches   80; Conservative   25; Mismatches 103; Indels 120; Ga

Qy   14 PLPRSLGSRDGGA----VRQAQPPAPMAARRSQRRRG------------------
               :||| |:|:|::| |||:||||
Db   6 PETESSGNRQISVAELLARQGVTGAP-ARRRRRRGDSDAITVAELTGEIPIIRDI

Qy   48 RGEPGTALLVPLALGLGLALACLGLLLAVVSLGSRASLSAQEPAQEELVAEE----
               : :| |||||||||||| |||||
Db   64 AGPDAHASQSPAANGR----------------VQVGRAAPQSPABFVAEQ-VAEEPTRT\

Qy   100 DQDPSELNPQTEESQDPAPFLNRLVRPRR-----SAPKGRKTRARRAI---AAHY--
               : :: | : :: :
Db   110 QPEPRWPKSPPQDRRESGPELSEYPRPLRHTHSDRAPAGPPSGAEHMSPDPVEHYPI

Qy   147 -----------EVHPRPGQDG---------AQAGVDGTVSGWEEARINSSS-----PI
                                        |||:
Db   170 DVLDTEVGEAEAETEVREAQPGRGERHAAAAAGTDVEGDGAAFAVARRALDVVP\

Qy   183 RQIGEFIVTR-------AGLYLYCQVHFDE----GKAVYLKLDL-----LVDGVI
               :| |||||:   ::| :::| :|| |
Db   230 ---GALVVLQSILAVAFGAGLF-----IAFDQLWRWNSIVALVLSVMVILGLVVSVF

Qy   226 CLEEFSAT------AASSLGPQLRLCQ 246
               :|| | || : : | ||
Db   282 KTEDIASTLIAVGALITLGP-LALLQ 308

RESULT 12
TN14_HUMAN
ID   TN14_HUMAN   STANDARD;   PRT;   240 AA.
AC   O43557; O75476; Q8WVF8; Q96LD2;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Tumor necrosis factor ligand superfamily member 14 (Herpesvirus er
DE   mediator-ligand) (HVEM-L).
GN   TNFSF14 OR LIGHT OR HVEML.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```
CC       NONHEMATOPOIETIC TUMOR LINES.
CC   -!- INDUCTION: UPREGULATED AFTER T-CELL ACTIVATION.
CC   -!- PTM: N-glycosylated.
CC   -!- PTM: The soluble form of isoform 1 derives from the membrane
CC       by proteolytic processing.
CC   -!- SIMILARITY: Belongs to the tumor necrosis factor family.
CC   -!- CAUTION: Ref.4 sequence differs from that shown due to a
CC       frameshift in position 178.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a coll
CC   between the Swiss Institute of Bioinformatics and the EMBL out
CC   the European Bioinformatics Institute. There are no restrictio
CC   use by non-profit institutions as long as its content is i
CC   modified and this statement is not removed. Usage by and for c
CC   entities requires a license agreement (See http://www.isb-sib.ch/
CC   or send an email to license@isb-sib.ch).

DR   EMBL; AF036581; AAC39963.1; -.
DR   EMBL; AF064090; AAC25169.1; -.
DR   EMBL; AY028261; AAK26160.1; -.
DR   EMBL; BC018058; AAH18058.1; ALT_FRAME.
DR   HSSP; P01375; 4TSV.
DR   Genew; HGNC:11930; TNFSF14.
DR   MIM; 604520; -.
DR   GO; GO:0005102; F:receptor binding; TAS.
DR   GO; GO:0006917; P:induction of apoptosis; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR006053; TNF_abc.
DR   InterPro; IPR006052; TNF_family.
DR   InterPro; IPR008983; TNF_like.
DR   InterPro; IPR003636; TNF_subf.
DR   Pfam; PF00229; TNF; 1.
DR   PRINTS; PR01234; TNECROSISFCT.
DR   ProDom; PD002012; TNF_subf; 1.
DR   SMART; SM002207; TNF; 1.
DR   PROSITE; PS00251; TNF_1; FALSE_NEG.
DR   PROSITE; PS00049; TNF_2; 1.
KW   Cytokine; Transmembrane; Glycoprotein; Signal-anchor;
KW   Alternative splicing.
FT   CHAIN         1    240   TUMOR NECROSIS FACTOR LIGAND SUPERF/
FT                           MEMBER 14, MEMBRANE FORM.
FT   CHAIN        ?83   240   TUMOR NECROSIS FACTOR LIGAND SUPERF/
FT                           MEMBER 14, SOLUBLE FORM.
FT   DOMAIN        1     37   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     38     58   SIGNAL-ANCHOR (TYPE-II MEMBRANE PRO.
FT                           (POTENTIAL).
FT   DOMAIN       59    240   EXTRACELLULAR (POTENTIAL).
FT   SITE         82     83   CLEAVAGE (POTENTIAL).
FT   DISULFID    154    187   POTENTIAL.
FT   CARBOHYD    102    102   N-LINKED (GLCNAC... .).
FT   VARSPLIC     38     73   Missing (in isoform 2).
FT                           /FTId=VSP 006452.
FT   CONFLICT    120    120   L -> V (IN REF. 4).
FT   CONFLICT    214    214   E -> K (IN REF. 2).
SQ   SEQUENCE   240 AA;  26351 MW;  49D0BF67E1390B39 CRC64;

Query Match             6.4%;  Score 92;  DB 1;  Length 240;
Best Local Similarity   23.7%;  Pred. No. 2.8;
Matches   44;  Conservative  19;  Mismatches  57;  Indels 66;  (

Qy   29  QAQPPAPMAARSQRRRGRRGEPGTALLVPLAIGLGLALACLGLLLAVVSLGSRA
             ||:  ||:||||:|:| : |:|||:::
Db   16  QTDIPFTRLGRSHRRQSCSVARVGLGLL-LLMGAGLAVQGWFLLQLHWRLG---

Qy   89  EPAQBELVAEEDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIA
                     |:||   :||:||
Db   67  ----EMV------TRLPDGPAGSWEQLIQERRS----------

Qy  149  HPRPGQDGAQGVDGTVSGWEEARINSSSPLRYNRQI-------GEFIVTH
             :||:  ||:|||   ::||       |       |||
Db   93  NPAAHLTGANSSLTG----------SGGPLLWETQLGLAFLRGLSYHDGALVTH

Qy  197  YLYCQV 202
```

heria; Primates; Catarrhini; Hominidae; Homo.
06;
[N.A. (ISOFORM 1).
340; PubMed=9462508;
bner R., Montgomery R.I., Kochel K.D., Cheung T.C.,
en S., Murphy M., Eisenberg R.J., Cohen G.H., Spear P.G.,

' member of the TNF superfamily, and lymphotoxin alpha are
erpesvirus entry mediator.";
-30(1998).

[N.A. (ISOFORM 1), AND CHARACTERIZATION.

532; PubMed=9765287;
McDonnell P.C., Brigham-Burke M., Lyn S.D., Minton J.,
le K., Spampanato J., Silverman C., Hensley P.,
Emery J.G., Deen K., Eichman C., Chabot-Fletcher M.,
ung P.R.;
entry mediator ligand (HVEM-L), a novel ligand for
mulates proliferation of T cells and inhibits HT29 cell

. 273:27548-27556(1998).

[N.A. (ISOFORM 2), AND PROCESSING.
948; PubMed=11673523;
Butrovich K.D., Houshmand P., Edwards W.R., Ware C.F.;
acterization of LIGHT reveals linkage to an immune
s on chromosome 19p13.3 and distinct isoforms generated
splicing or proteolysis.";
67:5122-5128(2001).

[N.A.

257; PubMed=12477932;
L., Feingold E.A., Grouse L.H., Derge J.G.,
Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
quellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
cEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
ton B., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
adan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Touchman J.W., Green E.D., Dickson M.C.,
Grimwood J., Schmutz J., Myers R.M.,
S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
Schein J.E., Jones S.J.M., Marra M.A.;
und initial analysis of more than 15,000 full-length
se cDNA sequences.";
cad. Sci. U.S.A. 99:16899-16903(2002).

[N.A.

Cytokine that binds to TNFRSF3/LTBR. Binding to the
eptor TNFRSF6B modulates its effects. Activates NFKB,
s the proliferation of T cells, and inhibits growth of
carcinoma HT-29. Acts as a receptor for Herpes simplex

Homotrimer.
AR LOCATION: Type II membrane protein and secreted
1), Cytoplasmic (isoform 2).
VE PRODUCTS:
ernative splicing; Named isoforms=2;

043557-1; Sequence=Displayed;
ynonyms=LIGHT delta-TM;
043557-2; Sequence=VSP 006452;
ECIFICITY: PREDOMINANTLY EXPRESSED IN THE SPLEEN BUT ALSO
THE BRAIN. WEAKLY EXPRESSED IN PERIPHERAL LYMPHOID
AND IN HEART, PLACENTA, LIVER, LUNG, APPENDIX, AND KIDNEY,
XPRESSION SEEN IN FETAL TISSUES, ENDOCRINE GLANDS, OR

```
|V 147

STANDARD;     PRT;   280 AA.

el. 41, Created)
el. 41, Last sequence update)
el. 41, Last annotation update)
factor ligand superfamily member 6 (FAS antigen ligand)
). OR CD95L.
quatus atys (Red-crowned mangabey) (Sooty mangabey).
azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
eria; Primates; Catarrhini; Cercopithecidae;
e; Cercocebus.
1;
N.A.
yes;
18; PubMed=11491535;
Bostik P.; Mayne A.E., King C.L., Genain C.P.,
sari A.A.;
encing, and homology analysis of nonhuman primate
and co-stimulatory molecules.";
53:315-328(2001).
Cytokine that binds to TNFRSF6/FAS, a receptor that
the apoptotic signal into cells. May be involved in
T cell mediated apoptosis and in T cell development.
S-mediated apoptosis may have a role in the induction of
tolerance, in the antigen-stimulated suicide of mature
r both. Binding to the decoy receptor TNFRSF6B/DcR3
its effects (By similarity).
onotrimer (Probable).
R LOCATION: Type II membrane protein and secreted (By
).
oluble form derives from the membrane form by
c processing (By similarity).
: Belongs to the tumor necrosis factor family.
--------------------------------------------------------
T entry is copyright. It is produced through a collaboration
wiss Institute of Bioinformatics and the EMBL outstation -
ioinformatics Institute. There are no restrictions on its
rofit institutions as long as its content is in no way
his statement is not removed. Usage by and for commercial
res a license agreement (See http://www.isb-sib.ch/announce/
il to license@isb-sib.ch).
--------------------------------------------------------
; AAK37606.1; -.
4TSV.
08064; Fas_ligand.
06053; TNF_abc.
06052; TNF_family.
08983; TNF_like.
03636; TNF_subf.
TNF; 1.
1; FASLIGAND.
4; TNECROSISFCT.
12; TNF_subf; 1.
; TNF; 1.
51; TNF_1; 1.
49; TNF_2; 1.
otosis; Transmembrane; Glycoprotein; Signal-anchor.
1     280    TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
             MEMBER 6, MEMBRANE FORM.
29    280    TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
             MEMBER 6, SOLUBLE FORM (BY SIMILARITY).
1     80     CYTOPLASMIC (POTENTIAL).
81    101    SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
             (POTENTIAL).
02    280    EXTRACELLULAR (POTENTIAL).
```

```
FT   DOMAIN       4     69       PRO-RICH.
FT   DOMAIN      45     64       POLY-PRO.
FT   SITE       128    129       CLEAVAGE (BY SIMILARITY).
FT   DISULFID   201    232       POTENTIAL.
FT   CARBOHYD   183    183       N-LINKED (GLCNAC. . .)  (POTENTIAL).
FT   CARBOHYD   249    249       N-LINKED (GLCNAC. . .)  (POTENTIAL).
FT   CARBOHYD   259    259       N-LINKED (GLCNAC. . .)  (POTENTIAL).
SQ   SEQUENCE   280 AA;  31407 MW;  729EA60067B7D398 CRC64;

Query Match             6.4%;  Score 92;  DB 1;  Length 280;
Best Local Similarity  20.8%;  Pred. No. 3.4;
Matches  62;  Conservative 44;  Mismatches 118;  Indels 74;  Ga

Qy   13 LPLPRSLGSRDGAVRQAQPPAPMAARRSQRRRGRRGBPGTALLVPLAL------
         ||||| ||                                  |||
Db   30 LPCPTSVPRRPGQRRPPPPPPPLP--------PPPPPLPPLPLPLPKKRGN

Qy   62 GLGLALACLGLLLAVVSLGSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQDPAF
         |||:||| :|:||| ::        :        :
Db   79 GLCLLVMFFMVLVALVGLG--LGMFQLFHLQKEL------AELRESTSQKHTASS

Qy  122 RLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEA-RINSS
         ::    ::  : :          :        :
Db  129 QIGHP---SPPPEKKEQRK--VAHLTGKPNSRSMPLE--------WEDTYGIVLLS

Qy  181 YNRQIGEFIVTRAGLYYLYCQVHFDEGKA---------VYLKLD----LLVDGVI
         ||||| |:    :  ||::
Db  175 YKK--GGLVINETGLYFVYSKVYF-RGQQSCTNLPLSHKVYMRNSKYPQDLVMMEGKN

Qy  226 CLEBFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFLTYFGLF(
         ||  :  :    ||:|:   :|:||
Db  232 CTTGQMWAHSSYLGAVFNLTSTDHLY------VNVSELSLVNFEESQ--TFFGLYt
```

RESULT 14

NFC4_HUMAN

```
ID   NFC4_HUMAN     STANDARD;     PRT;    902 AA.
AC   Q14934;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Nuclear factor of activated T-cells, cytoplasmic 4 (T cell
DE   transcription factor NFAT3) (NF-ATc4) (NF-AT3).
GN   NFATC4 OR NFAT3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=T-cell;
RX   MEDLINE=95269130; PubMed=7749981;
RA   Hoey T., Sun Y.-L., Williamson K., Xu X.;
RT   "Isolation of two new members of the NF-AT gene family and functi
RT   characterization of the NF-AT proteins.";
RL   Immunity 2:461-472(1995).
RN   [2]
RP   REVIEW.
RX   MEDLINE=99189746; PubMed=10089876;
RA   Crabtree G.R.;
RT   "Generic signals and specific outcomes: signaling through Ca2+,
RT   calcineurin, and NF-AT.";
RL   Cell 96:611-614(1999).
CC   -!- FUNCTION: Plays a role in the inducible expression of cytokine
CC       genes in T cells, especially in the induction of the IL-2 and
CC       4 (By similarity)
CC   -!- SUBUNIT: Member of the multicomponent NFATC transcription com
CC       that consists of at least two components, a pre-existing
CC       cytoplasmic component NFATC2 and an inducible nuclear componen
CC       NFATC1. Other members such as NFATC4, NFATC3 or members of th
CC       activating protein-1 family, MAF, GATA4 and Cbp/p300 can also
CC       the complex. NFATC proteins bind to DNA as monomers.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic for the phosphorylated form
```

RESULT 15

MLH1_MOUSE
```
ID   MLH1_MOUSE     STANDARD;      PRT;    760 AA.
AC   Q9JK91; O62454;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   DNA mismatch repair protein Mlh1 (MutL protein homolog 1).
GN   MLH1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RA   Kumaran M., Rao M.R.S.;
RT   "Cloning of the cDNA of the MutL homolog, MLH1 from mouse testis.
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 1-151 FROM N.A.
RX   MEDLINE=96270514; PubMed=8674118;
RA   Edelmann W., Cohen P.E., Kane M., Lau K., Morrow B., Bennett S.,
RA   Umar A., Kunkel T., Cattoretti G., Chaganti R., Pollard J.W.,
RA   Kolodner R.D., Kucherlapati R.;
RT   "Meiotic pachytene arrest in MLH1-deficient mice.";
RL   Cell 85:1125-1134(1996).
CC   -!- FUNCTION: Probably involved in the repair of mismatches in DN
CC   -!- SUBUNIT: HETERODIMER OF MLH1 AND PMS2 OR MLH1 AND MLH3 (BY
CC       SIMILARITY)
CC   -!- SIMILARITY: Belongs to the DNA mismatch repair mutL/hexB fami
CC
CC   This SWISS-PROT entry is copyright. It is produced through a coll
CC   between the Swiss Institute of Bioinformatics and the EMBL out
CC   the European Bioinformatics Institute. There are no restriction
CC   use by non-profit institutions as long as its content is in
CC   modified and this statement is not removed. Usage by and for c
CC   entities requires a license agreement (See http://www.isb-sib.ch/
CC   or send an email to license@isb-sib.ch).
DR   EMBL; AF250844; AAF64514.1; -.
DR   EMBL; U60872; AAC52672.1; JOINED.
DR   EMBL; U59881; AAC52672.1; JOINED.
DR   EMBL; U59882; AAC52672.1; JOINED.
DR   EMBL; U59883; AAC52672.1; JOINED.
DR   EMBL; U59884; AAC52672.1; JOINED.
DR   HSSP; P23367; 1BKN.
DR   MGD; MGI:101938; Mlh1.
DR   GO; GO:0000793; C:condensed chromosome; IDA.
DR   GO; GO:0007060; P:male meiosis chromosome segregation; IMP.
DR   GO; GO:0007126; P:meiosis; IDA.
DR   GO; GO:0007131; P:meiotic recombination; IMP.
DR   GO; GO:0006298; P:mismatch repair; IDA.
DR   InterPro; IPR003594; ATPbind_ATPase.
DR   InterPro; IPR002099; DNA_mis_repair;
DR   Pfam; PF01119; DNA_mis_repair; 1.
DR   Pfam; PF02518; HATPase_c; 1.
DR   SMART; SM00387; HATPase_c; 1.
DR   TIGRFAMs; TIGR00585; mutl; 1.
DR   PROSITE; PS00058; DNA_MISMATCH_REPAIR_1; 1.
KW   DNA repair.
SQ   SEQUENCE   760 AA;   84679 MW;   173C809372A29186 CRC64;

Query Match    6.3%;  Score 90.5;  DB 1;  Length 760;
Best Local Similarity   22.7%;  Pred. No. 14;
Matches    70;  Conservative   31;  Mismatches   126;  Indels   81;  G

Qy    13  LPLPRSLGSRDGGAVRQ-------AQPPAPMAARRSQRRRGRRGEPGTALLVPLA
              ||||||||||||||| :  :                   ::::
Db   436  LPAPAERAAAESENLERESLMETSDAAQKAAPTSSPGSSRKRHR--EDSDVEMVENA

Qy    65  LALAC-----LGLLAVVSLGSRASLSAQEPAQEELVAEEDQPSELNPQ--TEES
```

```
AAA79175.1; -.
462; -.
778; NFATC4.
-.
3; P:transcription co-activator activity; TAS.
4; P:inflammatory response; TAS.
6; P:transcription from Pol II promoter; TAS.
007110; IPT_TIG.
002909; IPT_TIG.
000451; NF_Rel_dor.
008366; NFAT
008967; P53-like.
; RHD; 1.
; TIG; 1.
89; NUCFACTORATC.
9; IPT; 1.
204; REL_1; FALSE_NEG.
254; REL_2; 1.
regulation; Activator; Nuclear protein; DNA-binding;
horylation.
62      69    POLY-PRO.
14     119    CALCINEURIN-BINDING.
13     293    2 APPROXIMATE SP REPEATS.
13     229    SP 1.
77     293    SP 2 (APPROXIMATE).
97     304    POLY-PRO.
68     270    NUCLEAR LOCALIZATION SIGNAL.
30     437    DNA-BINDING.
72     674    NUCLEAR LOCALIZATION SIGNAL.
2 AA;   95472 MW;   E59F15F7647A47C6 CRC64;

arity        6.4%;   Score 92;  DB 1;  Length 902;
           25.0%;  Pred. No. 13;
onservative   15;  Mismatches   73;  Indels   56;  Gaps   8;

SLGSRDGGAVRQAQPPAPMAARRS---QRRRGRRGEPGTALLVPLAIGLGHLALAC  69
            |:                        ||:
GRGPEDSWLLLSAPGPTPASPRPASPCGKRRYSSSGTPSSA--------SPALSR  286

LAVVSLGSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQDPAPFLNRLVRPRRS  129
               ||:
----SLGEGS----------EPPPPPL-PLARDPGSPGPFDYVGAPPAES  325

GRKTRARRAIA----------AHYEVHPRPGQDGAQAGVD----GTVSG  167
|:|:|:|                ||:|
TRTSSEQAVALPRSEEPASCNGKLPLGABESVAPPGGSRKEVAGMDYLAVPSPLA  385

RINSSSPL  179
||:
RIGGHSPI  397
```

```
ter activation that is controlled by calcineurin-
dephosphorylation. Rapid nuclear exit of NFATC is thought
mechanism by which cells distinguish between sustained
ient calcium signals. The subcellular localization of
y a key role in the gene transcription.
CIFICITY: Highly expressed in placenta, lung, kidney,
l ovary. Weakly expressed in spleen and thymus. Not
in peripheral blood lymphocytes.
al Similarity Domain (RSD) allows DNA-binding and
ve interactions with AP1 factors (By similarity).
phorylated by NFATC-kinase; dephosphorylated by
in (By similarity).
Y: Belongs to the Rel/Dorsal family.

OT entry is copyright. It is produced through a collaboration
Swiss Institute of Bioinformatics and the EMBL outstation -
Bioinformatics Institute. There are no restrictions on its
profit institutions as long as its content is in no way
this statement is not removed. Usage by and for commercial
ires a license agreement (See http://www.isb-sib.ch/announce/
ail to license@isb-sib.ch).
```

```
:YPRRRIINLTSVLSLQEEISERCHETLRE--ILRNHSFVGCVNPQWALAQHQTKL 551
        |  |:|:||         ||  : |  :  |          :|||    :|

RLVRPRRSAPKGRKTRARRAIAAHYEV---------HPRPGQDGAQAGVDG 163
         :  ::        |:|:                     :|

---------TTKLSEELFYQILIYDFANFGVLRLSEPAPLFDLAMLALDS 596

VEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYLKLDLLVDG--- 220
       |  ||| :   :  :  |   |||  :  :|  ||||:|  ||:||

VTEDDGPKEGLAEY---IVEFLKKKAEMLADYFSVEIDE-EGNLIGLPLLIDSYVP 652

--VLALR---------CLEEFSATAAS--SLGPQ--LRLCQVSGLLALRPGSS 258
   :  |           |  |       |     :|  : |||    |:||:

LPIFILRLATEVNWDEEKECPESLSKECAMFYSIRKQYILEESTLSGQQSDMPGST 712
```

```
CLPW 266
  ||
--PW 716
```

April 7, 2004, 17:45:19
  secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

search, using sw model

Jul 7, 2004, 17:37:32 ; Search time 56.9116 Seconds
                        (without alignments)
                        1409.967 Million cell updates/sec

09-245-198A-4

SLLDFEISARRLPLPRSLG.........PWAHLKAAPFLTYFGLFQVH 284

SUM62
Gap 10.0 , Gapext 0.5

6107 seqs, 282547505 residues

satisfying chosen parameters: 1586107

th: 0
th: 2000000000

Minimum Match 0%
Maximum Match 100%
listing first 45 summaries

Genseq_29Jan04:*
geneseqp1980s:*
geneseqp1990s:*
geneseqp2000s:*
geneseqp2001s:*
geneseqp2002s:*
geneseqp2003as:*
geneseqp2003bs:*
geneseqp2004s:*

the number of results predicted by chance to have a
than or equal to the score of the result being printed,
d by analysis of the total score distribution.

SUMMARIES

ry ch	Length	DB	ID		Description
.0	284	2	AAW47525	Aaw47525	Homo sapi
.8	249	2	AAY09369	Aay09369	Human tum
.8	249	2	AAY95338	Aay95338	Human PRO
.8	249	3	AAB07526	Aab07526	Amino aci
.8	249	3	AAU86129	Aau86129	Human PRO
.8	249	6	ABR42315	Abr42315	Human TWE
.8	249	7	ADC35206	Adc35206	Human TNF
.6	249	2	AAW29745	Aaw29745	TNF relat
.6	249	4	AAE00891	Aae00891	Human TRE
.6	249	7	ADC97712	Adc97712	Murine FL
.8	273	4	AAU03499	Aau03499	TWEAK ext
.5	208	4	AAW93590	Aaw93590	Human TNR
.6	225	2	AAW47524	Aaw47524	Mus muscu
.6	225	3	AAB07527	Aab07527	Amino aci
.5	211	2	AAW93591	Aaw93591	Mouse TNR
.8	189	2	AAW29746	Aaw29746	TNF relat
.8	189	4	AAE00892	Aae00892	Human UL4
.7	146	4	AAE00895	Aae00895	Human TRE
.0	325	4	ABB67553	Abb67553	Drosophil
.0	409	5	AAU77718	Aau77718	Drosophil
.5	211	3	AAY58216	Aay58216	Canine ma
.5	260	3	AAY58215	Aay58215	Canine CD
.4	254	2	AAR64190	Aar64190	Human 4-1
.4	254	2	AAM26657	Aam26657	Human 4-1
.4	254	5	ABB75953	Abb75953	Human cyt

26	106.5	7.4	254	6	ABR42312	Abr42312 Hum
27	106.5	7.4	254	7	ADC35200	Adc35200 Hum
28	106.5	7.4	254	7	ADD18780	Add18780 Hum
29	106	7.3	1428	3	AAY97033	Aay97033 Cas
30	105.5	7.3	406	5	AAU77717	Aau77717 Dro
31	104.5	7.2	779	5	ABB07845	Abb07845 Hum
32	104	7.2	409	5	AAU77716	Aau77716 Dro
33	104	7.2	1323	2	AAR55248	Aar55248 N-m
34	102	7.1	256	1	AAM25657	Aam25657 Hum
35	100.5	7.0	647	2	AAW04327	Aaw04327 Rat
36	100	6.9	220	4	AAB62340	Aab62340 Gpl
37	99	6.9	574	3	AAY97032	Aay97032 Cas
38	99	6.9	855	7	ADB96563	Adb96563 Hum
39	99	6.9	915	6	ABP56824	Abp56824 Hum
40	99	6.9	1008	4	AAM78891	Aam78891 Hum
41	99	6.9	1020	4	AAM79875	Aam79875 Hum
42	97.5	6.8	1560	7	ADE71264	Ade71264 Nov
43	96.5	6.7	1097	4	ABG25655	Abg25655 Nov
44	96.5	6.7	1631	4	ABG22481	Abg22481 Nov
45	96.5	6.7	19938	6	ABP76681	Abp76681 Str

ALIGNMENTS

RESULT 1
AAW47525
ID   AAW47525 standard; protein; 284 AA.
XX
AC   AAW47525;
XX
DT   21-JUL-1998 (first entry)
XX
DE   Homo sapiens tumour necrosis factor related ligand (TRELL).
XX
KW   TRELL; tumour necrosis factor related ligand; tnf; treatment; canc
KW   autoimmune disease; immune system; stimulation; suppression;
KW   graft rejection.
XX
OS   Homo sapiens.
XX
PN   WO9805783-A1.
XX
PD   12-FEB-1998.
XX
PF   07-AUG-1997;   97WO-US013945.
XX
PR   07-AUG-1996;   96US-0023541P.
PR   18-OCT-1996;   96US-0028515P.
PR   18-MAR-1997;   97US-0040820P.
XX
PA   (BIOJ) BIOGEN INC.
PA   (UYGE-) UNIV GENEVA FACULTY MEDICINE.
XX
PI   Chicheportiche Y, Browning JL;
XX
DR   WPI; 1998-145619/13.
DR   N-PSDB; AAV18600.
XX
PT   Tumour necrosis factor related ligand - useful for, e.g. treating
PT   auto-immune disease and immune responses to tissue grafts.
XX
PS   Claim 12; Page 50-51; 69pp; English.
XX
CC   The sequence is that of human tumour necrosis factor related ligar
CC   (TRELL). TRELL or active fragments can be included with a carrier
CC   pharmaceutical compositions to treat cancer, autoimmune diseases c
CC   immune responses to tissue grafts, or to stimulate or suppress the
CC   system. It is useful to screen for TRELL receptors, by labelling w
CC   detectable label and screening compositions for binding. Agents
CC   interfering with TRELL-receptor binding can also be screened for,
CC   then be administered, optionally with interferon-gamma, to induce
CC   death or treat, suppress or alter immune responses (especially inv

CC   lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has
CC   cytostatic activity. Apo-3 ligand can be used to induce apoptosis
CC   mammalian cancer cells, to induce NF-kappaB-dependent transcripti
CC   to induce JNK/SAPK-dependent responses in mammalian cells
XX
SQ   Sequence 249 AA;

Query Match   87.8%; Score 1268; DB 2; Length 249;
Best Local Similarity   100.0%; Pred. No. 1.3e-112;
Matches 249; Conservative   0; Mismatches   0; Indels   0; G

Qy   36 MAARRSQRRRGRGEPGTALLVPLALGLGLALACLGGLLLAVVSLGSRASLSAQEPA
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MAARRSQRRRGRGEPGTALLVPLALGLGLALACLGGLLLAVVSLGSRASLSAQEPA

Qy   96 VAEEDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPR
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VAEEDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPR

Qy   156 GAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYLYCQVHFDEGKAVY
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYLYCQVHFDEGKAVY

Qy   216 LLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGGSLRIRTLPWAHLKA
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 LLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGGSLRIRTLPWAHLKA

Qy   276 TYFGLFQVH 284
          |||||||||
Db   241 TYFGLFQVH 249

RESULT 3
AAY95338
ID   AAY95338 standard; protein; 249 AA.
XX
AC   AAY95338;
XX
DT   25-SEP-2000 (first entry)
XX
DE   Human PRO207 antitumour protein.
XX
KW   PRO207; human; antitumour; tumour; therapy; cytostatic; breast ca
KW   ovarian cancer; renal cancer; colorectal cancer; uterine cancer;
KW   prostate cancer; lung cancer; bladder cancer;
KW   central nervous system cancer; melanoma; leukaemia; neoplasm.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   Peptide   1..40
FT   /label= Signal_peptide
FT   Modified-site   10..14
FT   /note= "amidation"
FT   Peptide   24..35
FT   /note= "prokaryotic membrane lipoprotein lipid"
FT   Modified-site   27..33
FT   /note= "N-myristoylation"
FT   Modified-site   29..35
FT   /note= "N-myristoylation"
FT   Modified-site   36..42
FT   /note= "N-myristoylation"
FT   Protein   41..249
FT   /label= PRO207
FT   Modified-site   45..51
FT   /note= "N-myristoylation"
FT   Modified-site   97..101
FT   /note= "amidation"
FT   Modified-site   118..124
FT   /note= "N-myristoylation"
FT   Modified-site   121..127
FT   /note= "N-myristoylation"
FT   Modified-site   125..131
FT   /note= "N-myristoylation"

---

rcinoma cells) involving a signal pathway between TRELL and
It's coding sequence can be used in gene therapy for TRELL-
ders in mammals (especially humans), e.g. tumours,
d inflammatory diseases or inherited genetic disorders, by
nto cells, and expressing, therapeutically effective amounts
e.g. a virus comprising a gene encoding TRELL. It may also
the preparation of prepare probes for screening
etic DNAs for TRELL-encoding sequences and for antisense

AA;

arity   100.0%; Score 1444; DB 2; Length 284;
         100.0%; Pred. No. 2.4e-129;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   DFEISARRLPLPRSLGSRDGGAVRQAQPPAPMAARRSQRRRGRGEPGTALLVPLA 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DE   DFEISARRLPLPRSLGSRDGGAVRQAQPPAPMAARRSQRRRGRGEPGTALLVPLA 60

Qy   LALACLGLLAVVSLGSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQDPAPFL 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   LALACLGLLAVVSLGSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQDPAPFL 120

Qy   RPPRSAPKGRKTRARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLR 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   RPPRSAPKGRKTRARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLR 180

Qy   IGEFIVTRAGLYLYCQVHFDEGKAVYLKLDLLVDGVLALRCLEEFSATAASSLGP 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   IGEFIVTRAGLYLYCQVHFDEGKAVYLKLDLLVDGVLALRCLEEFSATAASSLGP 240

Qy   CQVSGLLALRPGGSLRIRTLPWAHLKRAAPFLTYFGLFQVH 284
        |||||||||||||||||||||||||||||||||||||||||
Db   CQVSGLLALRPGGSLRIRTLPWAHLKRAAPFLTYFGLFQVH 284

dard; protein; 249 AA.

(first entry)

necrosis factor Apo-3 ligand protein sequence.

necrosis factor; Apo-3 ligand; lymphotoxin; apoptosis;
pendent transcription; JNK/SAPK-dependent response; cancer.

NTECH INC.

Marsters SA, Pitti R;

7982/24.
5000.

3- ligand (a tumor necrosis factor) homologue.

1; 74pp; English.

sequence represents a human tumour necrosis factor (TNF) and

98WO-US021407.

97US-0062037P.
97US-0069862P.

/note= "N-myristoylation"
128..134
/note= "N-myristoylation"
139..143
/note= "Asn is N-glycosylated"

.

99WO-US028565.

98US-0113296P.
99WO-US005028.
99US-0130232P.
99US-0131445P.
99US-0134287P.
99US-0144758P.
99US-0145698P.
99WO-US021090.
99WO-US021547.

ECH INC.

Goddard A, Godowski PJ, Gurney AL, Marsters SA;
tti RM, Wood WI;

68/38.
17.

ion to inhibit neoplastic cell growth or for treating tumor
rises polypeptides PRO179, PRO207, PRO320, PRO219, PRO221,
, PRO301, PRO526, PRO362, PRO356, PRO509 or PRO866.

4; 172pp; English.

quence is that of human antitumour protein PRO207, as
foetal kidney cDNA clone (see AAA49717). PRO207 shows
uence identity to tumour necrosis factor family members,
an lymphotoxin-beta (23.4%) and human CD40 ligand (19.8%).
216. A claimed method for inhibiting the growth of a tumour
exposing the tumor cell to PRO179, PRO207, PRO320, PRO219,
, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or PRO866
49), their agonists or chimeric polypeptides incorporating
ur is especially a cancer selected from breast, ovarian,
tal, uterine, prostate, lung, bladder and central nervous
melanoma and leukaemia. Methods for the recombinant
the antitumour proteins are also provided

A;

rity   87.8%;   Score 1268;   DB 3;   Length 249;
nservative   100.0%;   Pred. No. 1.3e-112;
   0;   Mismatches   0;   Indels   0;   Gaps   0;

SQRRRGREPGTALLVPLALGLGLALACLGLLLAVVSLGSRASLSAQEPAQEEL 95
QDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQD 155
VDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYLKLD 215
VLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFL 275
FQVH 284

Db

RESULT 4
AAB07526
ID   AAB07526 standard; protein; 249 AA.
XX
AC   AAB07526;
XX
DT   20-OCT-2000 (first entry)
XX
DE   Amino acid sequence of a soluble recombinant human TWEAK protein.
XX
KW   TWEAK protein; immunological disorder; immune response; inflammati
KW   TWEAK blocking agent; autoimmune disease; organ transplant rejecti
KW   Graft-versus-Host disease; GVHD; lymphoid cell malignancy; shock;
XX
OS   Homo sapiens.
XX
XX   WO200042073-A1.
XX
PD   20-JUL-2000.
XX
PF   14-JAN-2000; 2000WO-US001044.
XX
PR   15-JAN-1999;   99US-0116168P.
XX
PA   (BIOJ) BIOGEN INC.
XX
PI   Rennert P;
XX
DR   WPI; 2000-476036/41.
XX
PT   Preventing and treating immune responses using modulators, especia
PT   antibodies, of TWEAK, TWEAK receptors and TWEAK ligands, useful fc
PT   treating e.g. inflammation and graft versus host disease.
XX
PS   Disclosure; Fig 1; 45pp; English.
XX
CC   The present sequence represents a TWEAK protein. The specificatior
CC   describes a method for preventing or treating an immunological dis
CC   and/or inhibiting an immune response in an animal. The method comr
CC   administering a TWEAK blocking agent. The method may be used for
CC   preventing and treating immune disorders associated with inappropi
CC   expression and/or activity of TWEAK. These disorders include autoi
CC   diseases, acute and chronic inflammation, organ transplant rejecti
CC   Graft-versus-Host disease (GVHD), lymphoid cell malignancies, sept
CC   other forms of shock, loss of immune responsiveness (as seen in hu
CC   immunodeficiency virus (HIV) infections) and failure of the immune
CC   response to tumour growth
XX
SQ   Sequence 249 AA;

Query Match   87.8%;   Score 1268;   DB 3;   Length 249;
Best Local Similarity   100.0%;   Pred. No. 1.3e-112;
Matches   249;   Conservative   0;   Mismatches   0;   Indels   0;   Ga

Qy   36   MAARRSQRRRGRRGEPGTALLVPLALGLGLALACLGLLLAVVSLGSRASLSAQEPAC
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1   MAARRSQRRRGRRGEPGTALLVPLALGLGLALACLGLLLAVVSLGSRASLSAQEPAC

Qy   96   VAEEDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRE
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61   VAEEDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRE

Qy   156   GAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYI
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121   GAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYI

Qy   216   LLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKAF
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181   LLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKAF

Query Match   87.8%; Score 1268; DB 5; Length 249;
Best Local Similarity 100.0%; Pred. No. 1.3e-112;
Matches 249; Conservative 0; Mismatches 0; Indels 0; G

Qy   36 MAARSQRRRGRRGEPGTALLVPLAIGLGLALACLGLLLAVVSLGSRASLSAQEPA
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MAARSQRRRGRRGEPGTALLVPLAIGLGLALACLGLLLAVVSLGSRASLSAQEPA

Qy   96 VAEEDQDPSELNPQTESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPR
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VAEEDQDPSELNPQTESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPR

Qy   156 GAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAVY
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAVY

Qy   216 LLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKA
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 LLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKA

Qy   276 TYFGLFQVH 284
         |||||||||
Db   241 TYFGLFQVH 249

RESULT 6
ABR42315
ID   ABR42315 standard; protein; 249 AA.
XX
AC   ABR42315;
XX
DT   11-AUG-2003 (first entry)
XX
DE   Human TWEAK protein.
XX
KW   Human; TWEAK; tumour necrosis factor; ligand; cytostatic;
KW   immunomodulator; osteopathic.
XX
OS   Homo sapiens.
XX
PN   WO2003040307-A2.
XX
PD   15-MAY-2003.
XX
PF   25-JUL-2002; 2002WO-US023782.
XX
PR   27-JUL-2001; 2001US-0307838P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Hilbert DH, Rosen CA;
XX
DR   WPI; 2003-430659/40.
DR   N-PSDB; ACC57901.
XX
PT   New heteromultimeric complex having a first polypeptide member of
PT   tumor necrosis factor (TNF) ligand family, and a second different
PT   of TNF ligand family, useful for treating cancer, osteoporosis or
PT   autoimmune disease.
XX
PS   Disclosure; Page 368-369; 388pp; English.
XX
CC   The present sequence is the protein sequence for human TWEAK prot
CC   invention relates to compositions comprising heterotrimeric compl
CC   tumour necrosis factor (TNF) ligand family members, and their use
CC   detection, prevention and treatment of disease. In one embodiment
CC   heterotrimeric complex comprises full-length or extracellular por
CC   TWEAK and full-length or extracellular portions of other TNF liga
CC   family members, preferably VEGI or VBGI-SV. The heterotrimeric co
CC   of the invention are useful for treating an autoimmune disease, c
CC   osteoporosis, and particularly for inhibiting cancer cell prolife
CC   increasing B cell proliferation, or inducing apoptosis of T cells
XX

```
\;
           87.8%;  Score 1268;  DB 6;  Length 249;
city 100.0%;  Pred. No. 1.3e-112;  Indels    0;  Gaps    0;
nservative    0;  Mismatches    0;  Indels    0;

SQRRGRRGEPGTALLVPLALGLGLALACLGLLLAVVSLGSRASLSAQEPAQEEL  95
||||||||||||||||||||||||||||||||||||||||||||||||||||||
SQRRGRRGEPGTALLVPLALGLGLALACLGLLLAVVSLGSRASLSAQEPAQEEL  60

QDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQD  155
||||||||||||||||||||||||||||||||||||||||||||||||||||||
QDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQD  120

VDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYLKLD  215
||||||||||||||||||||||||||||||||||||||||||||||||||||||
VDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYLKLD  180

VLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFL  275
||||||||||||||||||||||||||||||||||||||||||||||||||||||
VLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFL  240

FQVH  284
||||
FQVH  249
```

```
ard; protein; 249 AA.

first entry)

nd family member #12.

necrosis factor; TNF ligand; endokine alpha;
resorption disorder; osteoporosis; Paget's disease;
fication.

.1.

002US-00218547.

001US-0312542P.
001US-0330761P.

C A.
LLI B.

Rosen CA, Nardelli B;

072/66.
205.

alpha gene useful for preparing a composition for treating a
ated with excessive or insufficient bone resorption e.g.,
Paget's disease or arterial calcification.

SEQ ID NO 24; 145pp; English.

relates to an isolated nucleic acid molecule encoding a
is factor family ligand. A composition comprising the
ody or its fragment is used for treating an individual in
ased level of endokine alpha activity. The endokine alpha
resent in a heterotrimeric complex is used for treating an
```

```
CC   individual having a disorder associated with excessive bone resorp
CC   e.g. osteoporosis, Paget's disease or arterial calcification. Trea
CC   individual having a disorder associated with insufficient bone res
CC   comprises administering an endokine alpha antagonist, which is the
CC   antibody that binds specifically to endokine alpha polypeptide. Th
CC   present sequence represents the amino acid sequence of a tumour ne
CC   factor family ligand.
XX
SQ   Sequence 249 AA;

Query Match              87.8%;  Score 1268;  DB 7;  Length 249;
Best Local Similarity  100.0%;  Pred. No. 1.3e-112;
Matches  249;  Conservative    0;  Mismatches    0;  Indels    0;  Ga

Qy    36 MAARRSQRRGRRGEPGTALLVPLALGLGLALACLGLLLAVVSLGSRASLSAQEPAQ
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MAARRSQRRGRRGEPGTALLVPLALGLGLALACLGLLLAVVSLGSRASLSAQEPAQ

Qy    96 VAEEDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRF
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 VAEEDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRF

Qy   156 GAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYI
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYI

Qy   216 LLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKAF
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 LLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKAF

Qy   276 TYFGLFQVH  284
         |||||||||
Db   241 TYFGLFQVH  249
```

```
RESULT 8
AAW29745
ID   AAW29745 standard; protein; 249 AA.
XX
AC   AAW29745;
XX
DT   27-OCT-1998  (first entry)
XX
DE   TNF related endothelium proliferative agent protein.
XX
KW   TNF; endothelium proliferative agent; TREPA; wound healing; cance
KW   tissue grafting; vascularisation; apoptosis; autoimmune; birth co
XX
OS   Homo sapiens.
XX
PN   WO9835061-A2.
XX
PD   13-AUG-1998.
XX
PF   12-FEB-1998;    98WO-US002859.
XX
PR   12-FEB-1997;    97US-00798692.
PR   10-FEB-1998;    98US-00021706.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Wiley SR;
XX
DR   WPI; 1998-447255/38.
DR   N-PSDB; AAV47613.
XX
PT   Detecting nucleic acid encoding TREPA - useful for diagnosis and
PT   treatment of autoimmune disease, tumours and inflammation.
XX
PS   Claim 16; Page 123-4; 142pp; English.
XX
CC   The TNF-related endothelium proliferative agent (TREPA), or its
CC   activators or agonists, are used to treat a deficit of TREPA, e.g
```

XX
DR   WPI; 2001-280760/29.
DR   N-PSDB; AAD04350.
XX
PT   Inducing angiogenesis in mammal at desired sites for promoting wo
PT   healing, by administering soluble fragment of extracellular domai
PT   tumor necrosis factor related endothelium proliferative agent pro
XX
PS   Claim 1; Col 75-76; 53pp; English.
XX
CC   The present invention relates to extracellular signal molecules,
CC   particularly members of tumour necrosis factor (TNF) family molec
CC   designated as TREPA (TNF related endothelium proliferative agent)
CC   Soluble biologically active TREPA are used to treat TREPA-associa
CC   diseases, tumours or metastases. TREPA is used for inducing angic
CC   in human for promoting wound healing and for vascularising grafte
CC   for successful grafting and to promote tissue grafts. The present
CC   acid sequence is clone ID #690050 human TREPA
XX
SQ   Sequence 249 AA;

Query Match   87.6%;   Score 1265; DB 4;   Length 249;
Best Local Similarity   99.6%;   Pred. No. 2.5e-112;
Matches 248; Conservative   1; Mismatches   0;   Indels   0;   G

```
Qy   36  MAARRSQRRRGRRGEPGTALLVPLALGLGLALACLGLLLAVVSLGSRASLSAQEPA
             ||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MAARRSQKRRGRRGEPGTALLVPLALGLGLALACLGLLLAVVSLGSRASLSAQEPA

Qy   96  VAEEDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPF
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  VAEEDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPF

Qy  156  GAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAV
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  GAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAV

Qy  216  LLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLK
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  LLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLK

Qy  276  TYFGLFQVH 284
             |||||||||
Db  241  TYFGLFQVH 249
```

RESULT 10
ADC97712
ID   ADC97712 standard; protein; 249 AA.
XX
AC   ADC97712;
XX
DT   15-JAN-2004 (first entry)
XX
DE   Murine FL-TWEAK.
XX
KW   Murine; FL-TWEAK; TNF relatedness and weak ability to induce cel
KW   TNF; Tumour Necrosis Factor; TWEAK; fibrosis; cardiac disease;
KW   liver disease; lung disease; kidney disease; skin disease;
KW   skeletal muscle disease; adipose tissue disease;
KW   gastrointestinal tract disease; pancreatic disease;
KW   reproductive organ disease; neural disease; cartilage disease;
KW   bone disease; connective tissue disease; cellular death; hepatot
KW   dermatological; gastrointestinal; osteopathic.
XX
OS   Mus sp.
XX
PN   WO2003086311-A2.
XX
PD   23-OCT-2003.
XX
PF   09-APR-2003; 2003WO-US011350.
XX

---

healing or tissue grafting, by promoting vascularisation,
e apoptosis for treating cancer and eliminating autoreactive
n adjunct to cancer chemotherapy or antiviral treatment.
s can also be used to target cytotoxic agents or for
ation of the corresponding receptor, the nucleic acid for
used to transform tumour cells to render them more
TREPA and to screen for TREPA mimics. Ribozymes, antisense
ies or peptides, are used to treat TREPA-associated
. tumours and metastases (by inhibiting vascularisation),
or a wide range of autoimmune conditions, conditions
ormal stimulation of epithelial cells (e.g.
is), for birth control (inhibiting ovulation and placental
other angiogenic conditions (e.g. ulcers)

AA;

arity   87.6%;   Score 1265;   DB 2;   Length 249;
arity   99.6%;   Pred. No. 2.5e-112;
onservative   1; Mismatches   0;   Indels   0;   Gaps   0;

```
Qy   RSQRRRGRRGEPGTALLVPLALGLGLALACLGLLLAVVSLGSRASLSAQEPAQEEL  95
     ||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   RSQKRRGRRGEPGTALLVPLALGLGLALACLGLLLAVVSLGSRASLSAQEPAQEEL  60

Qy   DQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQD  155
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   DQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQD  120

Qy   GVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYLKLD  215
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   GVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYLKLD  180

Qy   GVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFL  275
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   GVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFL  240

Qy   LFQVH 284
     |||||
Db   LFQVH 249
```

dard; protein; 249 AA.

(first entry)

(TNF related endothelium proliferative agent).

: necrosis factor; TNF; angiogenesis; wound healing; TREPA;
endothelium proliferative agent; tumour; metastasis;
nerary.

Location/Qualifiers
98..249
/label= Extracellular_domain

98US-00105343.

97US-00798692.
98US-00021706.

TT LAB.

002US-0371611P.

```
KW   fusion protein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200145730-A2.
XX
PD   28-JUN-2001.
XX
PF   19-DEC-2000; 2000WO-US034755.
XX
PR   20-DEC-1999; 99US-0172878P.
PR   10-MAY-2000; 2000US-0203347P.
XX
PA   (IMMV ) IMMUNEX CORP.
XX
PI   Wiley SR;
XX
DR   WPI; 2001-417975/44.
DR   N-PSDB; AAS03964.
XX
PT   Modulating angiogenesis in a mammal for treating diseases mediate
PT   angiogenesis, e.g. solid tumors and vascular deficiencies of card
PT   peripheral tissue, by administering antagonist or agonist of TWEA
PT   receptor.
XX
PS   Example 1; Page 41; 46pp; English.
XX
CC   The sequence represents a a fusion protein encoded by the express
CC   vector pDC409-LZ-TWEAK. The fusion protein comprises a growth hor
CC   leader, a leucine zipper multimerisation domain, and the extracel
CC   domain of human TWEAK. The fusion protein was used in the isolati
CC   human TWEAK receptor (TWEAKR)-expressing clones from a COS cell h
CC   cDNA library. The TWEAK protein is a member of the tumour necrosi
CC   (TNF) family and induces angiogenesis. TWEAKR may therefore be us
CC   screen for and develop TWEAK agonists and antagonists for the mo
CC   of angiogenesis, to be used in the treatment and diagnosis of hum
CC   disease. The disorders mediated by angiogenesis include ocular di
CC   characterised by ocular neovascularisation such as diabetic retin
CC   neovascular glaucoma, retinoblastoma, retinopathy of prematurity,
CC   retrolental fibroplasia, rubeosis, uveitis, macular degeneration
CC   corneal graft neovascularisation, and inflammatory diseases such
CC   arthritis, rheumatism and psoriasis. Other treatable diseases inc
CC   malignant and metastatic conditions such as sarcomas and carcinom
CC   benign tumours and preneoplastic conditions, myocardial angiogene
CC   haemophilic joints, scleroderma, vascular adhesions, atherosclerc
CC   plaque neovascularisation, telangiectasia, wound granulation, cor
CC   atherosclerosis, peripheral atherosclerosis and ischaemia
XX
SQ   Sequence 273 AA;
```

Query Match   73.8%; Score 1066; DB 4; Length 273;
Best Local Similarity   100.0%; Pred. No. 3e-93;
Matches 207; Conservative   0; Mismatches   0; Indels   0; G

```
Qy    78  SLGSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKG
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    67  SLGSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKG

Qy   138  ARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRA
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   127  ARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRA

Qy   198  LYCQVHFDEGKAVYLKDLLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLAI
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   187  LYCQVHFDEGKAVYLKDLLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLAI

Qy   258  SLRIRTLPWAHLKAAPFLTYFGLFQVH  284
              ||||||||||||||||||||||||||||
Db   247  SLRIRTLPWAHLKAAPFLTYFGLFQVH  273
```

RESULT 12

```
Qy   197 YLYCQVHFDEGKAVYLKLDLLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLA
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 YLYCQVHFDEGKAVYLKLDLLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLA

Qy   257 SSLRIRTLPWAHLKAAPFLTYFGLFQVH 284
             |||||||||||||||||||||||||||
Db   181 SSLRIRTLPWAHLKAAPFLTYFGLFQVH 208
```

RESULT 13
```
AAW47524
ID   AAW47524 standard; protein; 225 AA.
XX
AC   AAW47524;
XX
DT   21-JUL-1998  (first entry)
XX
DE   Mus musculus tumour necrosis factor related ligand (TRELL).
XX
KW   TRELL; tumour necrosis factor related ligand; tnf; treatment; car
KW   abnormality; gestational abnormality; prostate cancer;
KW   autoimmune disease; immune system; stimulation; suppression;
KW   graft rejection.
XX
OS   Mus musculus.
XX
FH   Key            Location/Qualifiers
FT   Domain         1..21
FT                  /note "hydrophobic, transmembrane domain"
XX
PN   WO9805783-A1.
XX
PD   12-FEB-1998.
XX
PF   07-AUG-1997;  97WO-US013945.
XX
PR   07-AUG-1996;  96US-0023541P.
PR   18-OCT-1996;  96US-0028515P.
PR   18-MAR-1997;  97US-0040820P.
XX
PA   (BIOJ ) BIOGEN INC.
PA   (UYGE-) UNIV GENEVA FACULTY MEDICINE.
XX
PI   Chicheportiche Y,  Browning JL;
XX
DR   WPI; 1998-145619/13.
DR   N-PSDB; AAV18599.
XX
PT   Tumour necrosis factor related ligand - useful for, e.g. treatin
PT   auto-immune disease and immune responses to tissue grafts.
XX
PS   Claim 12; Page 48-50; 69pp; English.
XX
CC   The sequence is that of mouse tumour necrosis factor related lig
CC   (TRELL). TRELL or active fragments can be included with a carrie
CC   pharmaceutical compositions to treat cancer, autoimmune diseases
CC   immune responses to tissue grafts, or to stimulate or suppress t
CC   system. It is useful to screen for TRELL receptors, by labelling
CC   detectable label and screening compositions for binding. Agents
CC   interfering with TRELL-receptor binding can also be screened for
CC   then be administered, optionally with interferon- gamma, to indu
CC   death or treat, suppress or alter immune responses (especially i
CC   human adenocarcinoma cells) involving a signal pathway between T
CC   its receptor. It's coding sequence can be used in gene therapy f
CC   related disorders in mammals (especially humans), e.g. tumours,
CC   autoimmune and inflammatory diseases or inherited genetic disord
CC   introducing into cells, and expressing, therapeutically effectiv
CC   of a vector, e.g. a virus comprising a gene encoding TRELL. It n
CC   be of use in the preparation of prepare probes for screening
CC   natural/synthetic DNAs for TRELL-encoding sequences and for anti
CC   therapy
XX
SQ   Sequence 225 AA;
```

(left column, rotated)

```
dard; protein; 208 AA.


(first entry)


rotein.

is factor receptor; signal transducer molecule; TNF; APO4;
abnormality; gestational abnormality; prostate cancer;
PO9; TNRL-3; TNRL-1; diagnosis; treatment; therapy; disease;
omain; immunogen; antibody preparation; breast carcinoma;
man.




98WO-US018393.

97US-00924634.


WASHINGTON.



191/17.
424.


rosis Factor family receptor polypeptides and ligands -
agnosis and treatment of prostate cancer and developmental
l abnormalities.

] 13A; 156pp; English.

on describes isolated Tumor Necrosis Factor (TNF) family
/peptides: APO4, APO6, APO8 and APO9 or their active
d isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
fragments. APO4 is useful for diagnosing prostate cancer by
levels of APO4 in an individual. Prostate cancer can also be
g APO4 selective binding agents linked to a therapeutic
polypeptides are also useful for identifying selective
ts, useful in diagnosis/treatment of disease by binding of
e polypeptide/active fragment which is extracellular, or
the cell surface. The binding is preferably performed in
olypeptides/ active fragments are also useful for screening
and antagonists by binding and observing the changer in APO4
fective pharmacological agents useful in diagnosis or
disease are also identified using APO4 polypeptides/active
d APO4 signal transducer molecules that specifically interact
lasmic domain of APO4 and detecting a change in level of APO4
e method is performed in vivo or in vitro. APO polypeptides
ul as immunogens for preparing antibodies. APO4 is also
iagnosis/treatment of developmental or gestational
s. APO8 was transfected to human breast carcinoma cell line
nduced apoptosis

AA;

larity   73.5%;    Score 1062;  DB 2;   Length 208;
         99.5%;    Pred. No. 4.9e-93;
Conservative    0;  Mismatches    1;  Indels    0;   Gaps    0;

GSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKT 136
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKT 60

RATAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLY 196
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
RATAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYRQIGEFIVTRAGLY 120
```

Best Local Similarity 88.8%; Pred. No. 5.7e-89;
Matches 199; Conservative 9; Mismatches 16; Indels 0; G

```
Qy   61 LGLGLALACLGLLLAVVSLGSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQDP
           |||||||||||| :||||  :||||||| ||||:||| : :||||||||||||
Db    2 LSLGLALACLGLLLVVVSLGSWATLSAQEPSQEELTAEDRREPPELNPQTEESQDV

Qy  121 NRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSS
           :||||||||||||  ||||||||||||||||||||||||||||| : :||||
Db   62 EQLVRPRRSAPKGRKARPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSS

Qy  181 YNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYLKLDLILVDGVLALRCLEEFSATAAS
           |:||||| : :||||||||||||||||||||||||||: :||||||||||||||||
Db  122 YDRQIGEFTVIRAGLYYLYCQVHFDEGKAVYLKLDLLVNGVLALRCLEEFSATAAS

Qy  241 QLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 284
           |||||||||| ||||||||||||||||||||||||||||||||
Db  182 QLRLCQVSGLLPLRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 225
```

RESULT 15
AAW93591
ID   AAW93591 standard; protein; 211 AA.
XX
AC   AAW93591;
XX
DT   18-JUN-1999 (first entry)
XX
DE   Mouse TNRL3 protein.
XX
KW   Tumour necrosis factor receptor; signal transducer molecule; TNF;
KW   developmental abnormality; gestational abnormality; prostate ca
KW   APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy;
KW   cytoplasmic domain; immunogen; antibody preparation; breast carci
KW   apoptosis; mouse.
XX
OS   Mus sp.
XX
PN   WO9911791-A2.
XX
PD   11-MAR-1999.
XX
PF   04-SEP-1998;   98WO-US018393.
XX
PR   05-SEP-1997;   97US-0924634.
XX
PA   (UNIW) UNIV WASHINGTON.
XX
PI   Chaudhary PM;
XX
DR   WPI; 1999-205191/17.
DR   N-PSDB; AAX23425.
XX
PT   New Tumor Necrosis Factor family receptor polypeptides and ligand
PT   useful for diagnosis and treatment of prostate cancer and develop
PT   or gestational abnormalities.
XX
PS   Claim 40; Fig 13B; 156pp; English.
XX
CC   This invention describes isolated Tumor Necrosis Factor (TNF) fam
CC   receptor polypeptides: APO4, APO6, APO8 and APO9 (TNRL1 and T
CC   fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and T
CC   their active fragments. APO4 is useful for diagnosing prostate c
CC   determining levels of APO4 in an individual. Prostate cancer can
CC   treated using APO4 selective binding agents linked to a therapeut
CC   moiety. APO4 polypeptides are also useful for identifying select
CC   binding agents, useful in diagnosis/treatment of disease by bind
CC   agents to the polypeptide/active fragment which is extracellular,
CC   expressed on the cell surface. The binding is preferably perform
CC   vivo. APO4 polypeptides/ active fragments are also useful for sc
CC   for agonists and antagonists by binding and observing the change
CC   activity. Effective pharmacological agents useful in diagnosis o
CC   treatment of disease are also identified using APO4 polypeptides,

---

70.6%; Score 1020; DB 2; Length 225;
urity 88.8%; Pred. No. 5.7e-89;
nservative 9; Mismatches 16; Indels 0; Gaps 0;

```
ALACLGLLLAVVSLGSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQDPAPPL 120
|||||||| :||||  :||||||| ||||:||| : :||||||||
ALACLGLLLVVVSLGSWATLSAQEPSQEELTAEDRREPPELNPQTEESQDVVPFL 61

PRRSAPKGRKTRARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLR 180
|||||||||| ||||||||||||||||||||||||||||| : :||||
PRRSAPKGRKARPRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSSSPLR 121

GEFIVTRAGLYYLYCQVHFDEGKAVYLKLDLILVDGVLALRCLEEFSATAASSLGP 240
|| : :||||||||||||||||||||||||||: :||||||||||||||||
GEFTVIRAGLYYLYCQVHFDEGKAVYLKLDLLVNGVLALRCLEEFSATAASSPGP 181

QVSGLLALRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 284
|||| ||||||||||||||||||||||||||||||||
QVSGLLPLRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 225
```

lard; protein; 225 AA.

(first entry)

quence of a soluble recombinant murine TWEAK protein.

; immunological disorder; immune response; inflammation;
g agent; autoimmune disease; organ transplant rejection;
Host disease; GVHD; lymphoid cell malignancy; shock; tumour.

1.

2000WO-US001044.

99US-0116168P.

N INC.

036/41.

d treating immune responses using modulators, especially
f TWEAK, TWEAK receptors and TWEAK ligands, useful for
inflammation and graft versus host disease.

ig 1; 45pp; English.

equence represents a TWEAK protein. The specification
method for preventing or treating an immunological disorder
ting an immune response in an animal. The method comprises
a TWEAK blocking agent. The method may be used for
d treating immune disorders associated with inappropriate
d/or activity of TWEAK. These disorders include autoimmune
te and chronic inflammation, organ transplant rejection,
Host disease (GVHD), lymphoid cell malignancies, septic and
f shock, loss of immune responsiveness (as seen in human
ncy virus (HIV) infections) and failure of the immune
tumour growth

AA;

70.6%; Score 1020; DB 3; Length 225;

APO4 signal transducer molecules that specifically interact
asmic domain of APO4 and detecting a change in level of APO4
method is performed in vivo or in vitro. APO polypeptides
l as immunogens for preparing antibodies. APO4 is also
agnosis/treatment of developmental or gestational
. APO8 was transfected to human breast carcinoma cell line
duced apoptosis

AA;

arity   67.0%;   Score 968; DB 2; Length 211;
           89.1%; Pred. No. 4.8e-84;
onservative   9; Mismatches 14; Indels   0; Gaps   0;

```
/SLGSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKG 133
 ||| ||:|||||| |||||:  ::: :::||||||||||| :||||||||||||||
/SLGSWATLSAQEPSQEELTAEDRREPPELNPQTEESQDVVPFLEQLVRPRRSAPKG 60

?ARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRA 193
 :|||||||||||||||||||||||||||||||: :||||||||||:|||||:|:||
?PRRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEETKINSSSPLRYDRQIGEFTVIRA 120

?LYCQVHFDEGKAVYLKLDLLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLAL 253
 |||||||||||||||||||||:||||||||||||||||||| ||||||||||||||
?LYCQVHFDEGKAVYLKLDLLVNGVLALRCLEEFSATAASSPGPQLRLCQVSGLLAL 180

?SLRIRTLPWAHLKAAPFLTYFGLFQVH 284
 |||||||||||||||||||||||||||
?SLRIRTLPWAHLKAAPFLTYFGLFQVH 211
```

April 7, 2004, 17:44:47
secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

c search, using sw model

til 8, 2004, 19:08:51 ; Search time 5584 Seconds
(without alignments)
10657.232 Million cell updates/sec

-09-245-198A-3
73
tgtcattgttagacttga.........gacaaaatgttgataaatgg 1373

[GO_NUC
pop 60.0 , Gapext 60.0

70272 seqs, 2167516995 residues

es satisfying chosen parameters: 6940544

gth: 0
gth: 2000000000

isting first 100 summaries

enEmbl:*
gb_ba:*
gb_htg:*
gb_in:*
gb_om:*
gb_ov:*
gb_pat:*
gb_ph:*
gb_pl:*
gb_pr:*
gb_ro:*
gb_sts:*
gb_sy:*
gb_un:*
gb_vi:*
em_ba:*
em_fun:*
em_hum:*
em_in:*
em_mu:*
em_om:*
em_or:*
em_pat:*
em_ph:*
em_pl:*
em_ro:*
em_sts:*
em_un:*
em_vi:*
em_htg_hum:*
em_htg_inv:*
em_htg_other:*
em_htg_mus:*
em_htg_pln:*
em_htg_rod:*
em_htg_mam:*
em_htg_vrt:*
em_sy:*
em_htgo_hum:*
em_htgo_mus:*
em_htgo_other:*

score greater than or equal to the score of the result being pri
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1373	100.0	1373	6	BD062758	A.
2	1285	93.6	1306	9	AF030099	Hol
3	1172	85.4	1353	9	AX201324	Se
4	1172	85.4	1353	9	AY358870	Hol
5	1172	85.4	1368	9	AF055872	Hol
6	1172	85.4	1421	6	BD090952	Ap
7	958	69.8	1236	6	AR140407	Se
8	958	69.8	1236	6	BD057124	Me
9	796	58.0	1642	6	BC019047	Hol
10	625	45.5	898	6	AX180714	Se
11	620	45.2	60268	9	AC016876	Hol
12	553	40.3	1816	9	AY081051	Ho
13	435	31.7	218485	2	AC127470	Pa
14	179	13.0	195	6	AX379024	Se
15	145	10.6	180222	2	AC130192	Su
16	119	8.7	130254	2	AC136195	Ra
17	119	8.7	165316	2	AC119115	Ra
18	119	8.7	203083	2	AC069459	Mu
				2	AC098923	Ra
			3707		AC136563	Ra
			237		AL603707	M
			0100		AC126237	Ca
			921		AF030100	M
			.925		AC126921	Bo
			.239		AC126925	Ca
			.395		AC126239	Fe
			0954		AX201395	Se
			.9071		BD090954	Ap
					AC129071	Pa
29	46	3.4	1168	6	BD062757	A
30	31	2.3	234801	2	AC118309	Ra
31	26	1.9	26	6	AX201459	Se
32	26	1.9	140	6	AX522872	Se
33	26	1.9	145	6	AX522345	Se
34	26	1.9	145	6	AX522465	Se
35	26	1.9	153	6	AX322295	Se
36	26	1.9	281	6	AX209050	Se
37	26	1.9	353	6	BC062447	Hc
38	26	1.9	387	6	AR427020	Se
39	26	1.9	387	6	BD122573	Es
40	26	1.9	391	6	AX198506	Se
41	26	1.9	416	6	AR427022	Se
42	26	1.9	416	6	BD122575	Es
43	26	1.9	418	9	BC032480	Hc
44	26	1.9	439	6	AX330518	Se
45	26	1.9	439	6	AX410097	Se
46	26	1.9	452	6	AX895134	Se
47	26	1.9	483	6	AX381620	Se
48	26	1.9	486	6	BD224095	Se
49	26	1.9	487	6	AX408082	Se
50	26	1.9	495	9	BC060315	Hc
51	26	1.9	499	6	AR413525	Se
53	26	1.9	499	6	BD109078	Es
54	26	1.9	516	6	AX885407	Se
55	26	1.9	516	6	BD025017	Se
56	26	1.9	522	6	AR415934	Se
57	26	1.9	522	6	BD111487	Es
58	26	1.9	531	6	AX381939	Se
59	26	1.9	534	6	BC062222	Hc
60	26	1.9	538	6	AR415931	Se
61	26	1.9	538	6	BD111484	Es
62	26	1.9	540	6	AX885406	Se
63	26	1.9	540	6	BD025016	Se
64	26	1.9	540	6	BD224094	Se
65	26	1.9	540	6		

the number of results predicted by chance to have a

ORIGIN   /db_xref="taxon:32644"

Query Match   100.0%;   Score 1373;   DB 6;   Length 1373;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1373; Conservative 0; Mismatches 0; Indels 0;

```
Qy    1 ATGTCATTGTTAGACTTTGAAATTTCCGCCCGCGGCTTCCCCTTCCCCCGATCCCT
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGTCATTGTTAGACTTTGAAATTTCCGCCCGCGGCTTCCCCTTCCCCCGATCCCT

Qy   61 TCCCGGATGGGGGGCGGTGAGGCAGGCACAGCCCCCGCCCCCATGCCGCGCCCC
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TCCCGGATGGGGGGCGGTGAGGCAGGCACAGCCCCCGCCCCCATGCCGCGCCCC

Qy  121 AGCCAGAGGCGGAGGGGGCGCCGGGGGAGCCGGCACCGCCCTGCTCGTCGCCGCT
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AGCCAGAGGCGGAGGGGGCGCCGGGGGAGCCGGCACCGCCCTGCTCGTCGCCGCT

Qy  181 CTGGGCCTGGCCTGGCCTGCCCGCCCCAGGAGCCTGCCTCGGCCTCCTGCTCGTC
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CTGGGCCTGGCCTGGCCTGCCCGCCCCAGGAGCCTGCCTCGGCCTCCTGCTCGTC

Qy  241 AGCCGGGCATCGCTGTCCGCCCCAGGAGCCTGCCCAGGAGGAGAGTGGTGGCAGAGG
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 AGCCGGGCATCGCTGTCCGCCCCAGGAGCCTGCCCAGGAGGAGAGTGGTGGCAGAGG

Qy  301 CAGGACCCGTCGAACTGAATCCCAGACAGAGAAGCCAGAATCCTGCGCCTT
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CAGGACCCGTCGAACTGAATCCCAGACAGAGAAGCCAGAATCCTGCGCCTT

Qy  361 AACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGAAAACACGGGCTC
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 AACCGACTAGTTCGGCCTCGCAGAAGTGCACCTAAAGGCCGAAAACACGGGCTC

Qy  421 GCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACGGAGGCGCAGG
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACGGAGGCGCAGG

Qy  481 GTGGACGGACAGTGAGTGCTGGGGAGAAGCCAGAATCAACAGCTCCAGCCCTC
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 GTGGACGGACAGTGAGTGCTGGGGAGAAGCCAGAATCAACAGCTCCAGCCCTC

Qy  541 TACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTACTACTGT
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 TACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTACTACTGT

Qy  601 CAGGTGCACTTTGATGAGGGAAGGCCCTGTCTCCCACAGTCGTCCGAGCTGCGGCT
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 CAGGTGCACTTTGATGAGGGAAGGCCCTGTCTCCCACAGTCGTCCGAGCTGCGGCT

Qy  661 GTGCTGGCCCTGCCCTGCTGGGCCTGGGAATTCTCAGCCACTCGCGGCCAGTTCCCTCG
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 GTGCTGGCCCTGCCCTGCTGGGCCTGGGAATTCTCAGCCACTCGCGGCCAGTTCCCTCG

Qy  721 CAGCTTCCGCCTGCCAGGTGTCTGGCCCATCTCAAGGCTGCCCTGCGGGTCCAGGGTCTCCC
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 CAGCTTCCGCCTGCCAGGTGTCTGGCCCATCTCAAGGCTGCCCTGCGGGTCCAGGGTCTCCC

Qy  781 ATCCGCACCCTCCGTGGGGCCCTGGTCTCCCACCTCAAGGCTGCCCTGCGGGTCCAGGGTCTCCC
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 ATCCGCACCCTCCGTGGGGCCCTGGTCTCCCACCTCAAGGCTGCCCTGCGGGTCCAGGGTCTCCC

Qy  841 TTCCAGGTTCACTGAGGGGCCCTGGTCCCACAGTCGTCCGACAGTCGTCCCAGCGTGCCGGCT
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 TTCCAGGTTCACTGAGGGGCCCTGGTCCCACAGTCGTCCGACAGTCGTCCCAGCGTGCCGGCT

Qy  901 CGACAGCTCTCTGGGCACCCGGTCCCCTGCCCACCCTGCGCCACCCTGCAGCCGGCTCTTTGCT
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 CGACAGCTCTCTGGGCACCCGGTCCCCTGCCCACCCTGCGCCACCCTGCAGCCGGCTCTTTGCT

Qy  961 CCTGCCCCTCCCTCCTCTAGAGGCTGCCTGGGCCTTGGCTGTTTCATCCCACA
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
```

ALIGNMENTS

```
1.9   555    6   AR415935   AR415935   Sequence
1.9   555    6   BD11488    BD11488    EST and e
1.9   560    9   BC042807   BC042807   Homo sapi
1.9   595    6   AR415932   AR415932   Sequence
1.9   595    6   BD11485    BD11485    EST and e
1.9   1822   9   AK130060   AK130060   Homo sapi
1.9   2183   9   HSM02687   AL359585   Homo sapi
1.9   2650   9   AK124175   AK124175   Homo sapi
1.9   110293 9   HSDJ686N3  AL049766   Human DNA
1.9   159020 2   AC027023   AC027023   Homo sapi
1.8   507    6   AR427029   AR427029   Sequence
1.8   507    6   BD122582   BD122582   EST and e
1.8   40     6   BD090955   BD090955   Apo-3 lig
1.7   478    6   AR427041   AR427041   Sequence
1.7   478    6   BD122594   BD122594   EST and e
1.7   495    6   AR427028   AR427028   Sequence
1.7   495    6   BD122581   BD122581   EST and e
1.7   23         AX201458   AX201458   Sequence
1.7   164117 5   AL845513   AL845513   Zebrafish
1.7   172618 2   AC144822   AC144822   Danio rer
1.7   214279 2   AC138245   AC138245   Mus muscu
1.7   248178 2   AC093996   AC093996   Rattus no
1.6   38     6   BD090956   BD090956   Apo-3 lig
1.6   1005   6   AX695886   AX695886   Sequence
1.6   1173   1   AF182516   AF182516   Pseudomon
1.6   2942   6   AX695885   AX695885   Sequence
1.6   2942  10   MMU250723  AJ250723   Mus muscu
1.6   3642  10   BC046524   BC046524   Mus muscu
1.6   10989 10   AF45013854 AF450141   Mus muscu
1.6   41037  2   AC017476   AC017476   Drosophil
1.6   50295  6   AX695884   AX695884   Sequence
1.6   99416  9   AC105276   AC105276   Homo sapi
1.6   110000 2   AC110828_4 Continuation (5 of
1.6   110000 2   AC110828_5 Continuation (6 of
1.6   114252 5   BX649441   BX649441   Zebrafish
```

758   1373 bp   DNA   linear   PAT   27-AUG-2002
or necrosis factor related ligand.
758.1 GI:22608361
01505407-A/2.
ntified
ntified
ssified
ases 1 to 1373)
eportiche,Y. and Browning,J.L.
or necrosis factor related ligand
t: JP 2001505407-A 2 24-APR-2001;
N INC,THE FACULTY OF MEDICINE OF THE UNIVERSITY OF GENEVA
TNF family related protein
JP 200150540T-A/2
24-APR-2001
07-AUG-1997 JP 1998508239
07-AUG-1996 US   60/023541,18-OCT-1996 US   60/028515 PR
R-1997 US   60/040820
YVES CHICHEPORTICHE,JEFFREY L BROWNING
C12N15/28,C07K14/525,G01N33/68,C07K16/24,C12N15/11,A61K48/00,
C12N5/10,
A61K39/395,A61K38/19,C07K14/705,C12N15/12
Strandedness: Double;
Topology: Linear;
Key   Location/Qualifiers
CDS   1..852.
             Location/Qualifiers
1..1373
/organism="unidentified"
/mol_type="genomic DNA"

```
Query Match          93.6%;   Score 1285;  DB 9;   Length 1306;
Best Local Similarity 100.0%;  Pred. No. 0;
Matches 1285;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy   89 CACAGCCCCCGCCCCCCATGGCCGCCGTCGGAGCCAGGCGGCAGAGGCGGAGGGGGCGCCGG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 CACAGCCCCCGCCCCCCATGGCCGCCGTCGGAGCCAGGCGGCAGAGGCGGAGGGGGCGCCGG

Qy  149 AGCCGGCACCGCCCCTGCTCGTGTCCGTCGCCTGGGCTCGGCCTGGGCCTGGGCCTGGCC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AGCCGGCACCGCCCCTGCTCGTGTCCGTCGCCTGGGCTCGGCCTGGGCCTGGGCCTGGCC

Qy  209 TCGGCCTCCTGCTGCCGTGCCGTGGTCAGTTGGGGAGGAGTTGGGGAGGAGTCGCGCCAG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 TCGGCCTCCTGCTGCCGTGCCGTGGTCAGTTGGGGAGGAGTTGGGGAGGAGTCGCGCCAG

Qy  269 CTGCCCAGGAGAGTGGTGGCGCAGAGGAGGAGGACCAGGACCCGTCGAACTGAATCCC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CTGCCCAGGAGAGTGGTGGCGCAGAGGAGGAGGACCAGGACCCGTCGAACTGAATCCC

Qy  329 CAGAGAAAGCCAGGATCCTGCGCCCTTCCTGAACCGACTAGTTCGGCCTTCGCAGA
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 CAGAGAAAGCCAGGATCCTGCGCCCTTCCTGAACCGACTAGTTCGGCCTTCGCAGA

Qy  389 CACCTAAAGGCCGGAAAAACACGGGCTCGAAGAGCAGCCCAGCCATTATGAAGTT
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CACCTAAAGGCCGGAAAAACACGGGCTCGAAGAGCAGCCCAGCCCATTATGAAGTT

Qy  449 CACGACCTGGACAGGACGAGCGCAGGCAGGCAGGTGTGGACGGACAGTGAGTGGCTGG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 CACGACCTGGACAGGACGAGCGCAGGCAGGCAGGTGTGGACGGACAGTGAGTGGCTGG

Qy  509 AAGCCAGAATCAACAGCTCCAGCCCTCGCGCTACACCGGCTACAACCGCCAGATCGGGGAGTTT
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 AAGCCAGAATCAACAGCTCCAGCCCTCGCGCTACACCGGCTACAACCGCCAGATCGGGGAGTTT

Qy  569 TCACCCGGCTGGCTCTACTACCTGTACTGTCAGGTGCACTTTGATGAGGGGAAG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 TCACCCGGCTGGCTCTACTACCTGTACTGTCAGGTGCACTTTGATGAGGGGAAG

Qy  629 TCTACCTGAAGCTGGACTTGGTGGTGGATGGTGTGCTGGCCTGCCTGCTGGAG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 TCTACCTGAAGCTGGACTTGGTGGTGGATGGTGTGCTGGCCTGCCTGCTGGAG

Qy  689 TCTCAGCCACTGCGGCCAGTTCCCTGCGCCCTCGCGCCTCTGCCAGGTGTCT
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 TCTCAGCCACTGCGGCCAGTTCCCTGCGCCCTCGCGCCTCTGCCAGGTGTCT

Qy  749 TGTTGGCCCTGCGGCCAGGGTCCTCCCTGCGGATCGGCACCCTCCCTGGGCCCAT
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 TGTTGGCCCTGCGGCCAGGGTCCTCCCTGCGGATCGGCACCCTCCCTGGGCCCAT

Qy  809 AGGCTGCCCCCCTTCCCACTACTTCGGACTCTTCCAGGTTCACTGAGGGGCCCTC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 AGGCTGCCCCCCTTCCCACTACTTCGGACTCTTCCAGGTTCACTGAGGGGCCCTC

Qy  869 CCCCACAGTCGTCCCAGGCTGCCGGCCTCCCCTCGACAGCTCTCTGGGCACCCGGTC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 CCCCACAGTCGTCCCAGGCTGCCGGCCTCCCCTCGACAGCTCTCTGGGCACCCGGTC

Qy  929 CTGCCCACCCCAGCGCCGCTCTTGCTCCAGACCTGCCCTCCTCTAGAGGCTGC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 CTGCCCACCCCAGCGCCGCTCTTGCTCCAGACCTGCCCTCCTCTAGAGGCTGC

Qy  989 GCCTGTTCACGTGTTTTCCATCCCACATAAATACAGTATTCCCACTCTTATCTTAC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 GCCTGTTCACGTGTTTTCCATCCCACATAAATACAGTATTCCCACTCTTATCTTAC

Qy 1049 CCCCACACGCCCACTCTCCACCTCACTAGCTCCCAATCCTGACCCTTTGAGCC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 CCCCACACGCCCACTCTCCACCTCACTAGCTCCCAATCCTGACCCTTTGAGCC

Qy 1109 GTGATCTCGACTCCCCCGGCCACAGACCCCCAGGGCATTGTTGTTCACTGTACTC
```

Right column continuation:
```
...CCCCTCCCTGAGAGGCTGCTGGGCTGTTCACGTGTTTTCCATCCCACATAAAT 1020
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
...ATTCCCACTCTTATCTTACAACTCCCCACCGCCCACCCTCTCCACCTCACTAGCTC 1080
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
...ATTCCCACTCTTATCTTACAACTCCCCACCGCCCACCCTCTCCACCTCACTAGCTC 1080

...TCCCTGACCCTTTGAGCCCCCCAGTGATCTGACTCCCCCCTGGCCACAGACCCCC 1140
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
...TCCCTGACCCTTTGAGCCCCCCAGTGATCTGACTCCCCCCTGGCCACAGACCCCC 1140

...GCATTGTTGTTCACTGTACTCTGTGGCAAGGATGGGTCAAGAGACCCCACTTCAG 1200
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
...GCATTGTTGTTCACTGTACTCTGTGGCAAGGATGGGTCAAGAGACCCCACTTCAG 1200

...CAAGAGGGGCTGGACCTGGGCGGCGGCAGGAAGCCAAAAGAGACTGGGCCTCAGGAG 1260
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
...CAAGAGGGGCTGGACCTGGGCGGCGGCAGGAAGCCAAAAGAGACTGGGCCTCAGGAG 1260

...CAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGAT 1320
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
...CAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGAT 1320

...AAAACAGATATTATTTTTATTATTATTGTGACAAAATGTTGATAAATGG 1373
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
...AAAACAGATATTATTTTTATTATTATTGTGACAAAATGTTGATAAATGG 1373
```

```
99                          1306 bp    mRNA       linear   PRI 20-DEC-1997
apiens TWEAK mRNA, complete cds.
99
99.1  GI:2707218
apiens (human)
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 1306)
portiche,Y., Bourdon,P.R., Xu,H., Hsu,Y.M., Scott,H.,
n,C., Garcia,I. and Browning,J.L.
a new secreted ligand in the tumor necrosis factor family
eakly induces apoptosis
1. Chem. 272 (51), 32401-32410 (1997)
15
9
ses 1 to 1306)
a,P., Hession,C., Tizard,R. and Browning,J.
Submission
ted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center,
dge, MA 02142, USA
Location/Qualifiers
1..1306
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /chromosome="17"
        /map="17p13"
        /tissue_type="tonsil"
18..767
        /note="ligand in the TNF family; secreted protein; start
        codon not verified experimentally"
        /codon_start=1
        /product="TWEAK"
        /protein_id="AAC51923.1"
        /db_xref="GI:2707219"
        /translation="MAARRSQRRRGRRGEPGTALLVPLAIGLGLALACLGLLLAVVSL
        GSRASLSAQEPAQEELVAEEDQDPSELNPGTEESQDPAPFLNRLVRPRRSAPKGRKTR
        ARRAIAAHYEVHPRPGQDGAQAVDGTVSGWEEARINSSSPLRNRQIGEFIVTRAGL
        YYLYCQVHFDEGKAVYLKLDLLVDGVLALRCLEEFSATRASSLGPQLRLCQVSGLLAL
        RPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH"
```

Qy 409 CGGGCTCGAAGAGCGGATCGCAGCCCATTATGAGTTCATCCAGACCCTGACAGG
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 CGGGCTCGAAGAGCGGATCGCAGCCCATTATGAGTTCATCCAGACCCTGACAGG

Qy 469 GGCGAGGCAGTGTGGACGGACAGTGAGTTGGCTGGGAGGAAGCCAGAATCAACA
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 GGCGAGGCAGTGTGGACGGACAGTGAGTTGGCTGGGAGGAAGCCAGAATCAACA

Qy 529 AGCCCTCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGC
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 AGCCCTCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGC

Qy 589 TACCTGTACTGTCAGGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGG
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 TACCTGTACTGTCAGGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGG

Qy 649 CTGGTGGATGGTGTGCTGGCCCTGCCTGCCTGGAGGAATTCTCAGCCACTGCGGG
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 601 CTGGTGGATGGTGTGCTGGCCCTGCCTGCCTGGAGGAATTCTCAGCCACTGCGGG

Qy 709 TCCCTTCGGGCCCAGCTCCGCCTCTGCCAGGTGTCTGGGCTGTTGGCCCTGCGGGC
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 661 TCCCTTCGGGCCCAGCTCCGCCTCTGCCAGGTGTCTGGGCTGTTGGCCCTGCGGGC

Qy 769 TCCTCCCCTGCGGATCCGCACCCTCCCCTGGGCCCATTCTCAAGGCTGCCCCCTTCC
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 721 TCCTCCCTGCGGATCCGCACCCTCCCCTGGGCCCATTCTCAAGGCTGCCCCCTTCC

Qy 829 TACTTCGGACTCTTCCAGTTCACTGGGCACCCGGTCCCCTGCCCCACCCTCAG
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 781 TACTTCGGACTCTTCCAGTTCACTGGGCACCCGGTCCCCTGCCCCACCCTCAG

Qy 889 GCCGGCTCCCCTGACAGCTCTCTGGGCACCCCGGTCCCCGGTCCCGGCCTCCCACCCTCAG
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db 841 GCCGGCTCCCCTGACAGCTCTCTGGGCACCCCGGTCCCCGGTCCCGGCCTCCCACCCTCAG

Qy 949 CTTTGTCCAGACCTGCCCCTCCCCTCCCCTCTAGAGGCTGCCTGGCCTGTTCAGTGTT
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db 901 CTTTGTCCAGACCTGCCCCTCCCCTCCCCTCTAGAGGCTGCCTGGCCTGTTCAGTGTT

Qy 1009 TCCCACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCCACT
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 961 TCCCACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCCACT

Qy 1069 CCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCCC
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 CCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCCC

Qy 1129 GGCCAGAGCCCCCAGGGCATTGGTGTTCACTGTACTCGTGGGGCAAGGATGGGTCC
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 GGCCACAGACCCCCAGGGTCATTGTGTTCACTGTACTCGTGTGGGCAAGGATGGGTCC

Qy 1189 ACCCCACTTCAGGCACTAAGAGGGGCCTGGACCTGGGCGGCAGGAAGGCCAAAGAGAC
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 ACCCCACTTCAGGCACTAAGAGGGGCCTGGACCTGGGCGGCAGGAAGGCCAAAGAGAC

Qy 1249 CTAGGCCAGAGTTCCCAAATGTGAGGGGCCGAGAAACAAGACAAGCTCCTCCCTT
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 CTAGGCCAGGAGTTCCCAAATGTGAGGGGCCGAGAAACAAGACAAGCTCCTCCCTT

Qy 1309 TTCCCTGTGGATTTTTAAAACAGATATTATTTTATTATTGTGACAAAATGT
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 TTCCCTGTGGATTTTTAAAACAGATATTATTTTTATTATTGTGACAAAATGT

Qy 1369 AATGG 1373
            |||||
Db 1321 AATGG 1325

RESULT 4
AY358870
LOCUS   AY358870   1353 bp   mRNA   linear   PRI
DEFINITION Homo sapiens clone DNA30879 Apo-3 Ligand (UNQ181) mRNA,

---

(column 1 — left portion)

             |||||||||||||||||||||||||||||||||||
            TCTCGACTCCCCTGGCCACAGACCCCCCAGGGCATTGTTGTTCACTGTACTCTGTG 1080

            |||||||||||||||||||||||||||||||||||
AAGGATGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGCCGCA 1228
            |||||||||||||||||||||||||||||||||||
AAGGATGGGTCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGCCGCA 1140

            |||||||||||||||||||||||||||||||||||
AGCCAAAGAGACTGGGCCTAGGCCAGGAGTCCCAAATGTGAGGGGCGAGAAACAAG 1288
            |||||||||||||||||||||||||||||||||||
AGCCAAAGAGACTGGGCCTAGGCCAGGAGTCCCAAATGTGAGGGGCGAGAAACAAG 1200

            |||||||||||||||||||||||||||||||||||
AGCTCCTCCCCTTGAGAATTCCTGTGGATTTTTAAAACAGATATTATTTTTATTATT 1348
            |||||||||||||||||||||||||||||||||||
AGCTCCTCCCCTTGAGAATTCCTGTGGATTTTTAAAACAGATATTATTTTTATTATT 1260

            |||||||||||||||||||||||||||||||||||
GTGACAAAATGTTGATAAATGG 1373
            ||||||||||||||||||
GTGACAAAATGTTGATAAATGG 1285

324   1353 bp   DNA   linear   PAT 30-AUG-2001
nce 3 from Patent WO0153486.
324
324.1   GI:15391154

sapiens (human)
sapiens
yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
lia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

azi,A.J., Goddard,A., Godowski,P.J., Gurney,A.L.,
a,K.J., Marsters,S.A., Pan,J., Pitti,R.M., Roy,M.A., Smith,V.,
,D.M., Watanabe,C.K. and Wood,W.I.
sitions and methods for the treatment of tumour
t: WO 0153486-A 3 26-JUL-2001;
tech, Inc. (US)
Location/Qualifiers
1..1353
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

larity   85.4%;   Score 1172;   DB 6;   Length 1353;
Conservative 99.8%;   Pred. No. 0;   Mismatches   3;   Indels   0;   Gaps   0;

TCCCTCGGGTCCCGGGATGGGGGGCGGGGGGCGGTGAGGCAGCGGCACAGCCCCCGCCCCCATG 108
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
TCCCTCGGGTCCCGGGATGGGGGGCGGGGGGCGGTGAGGCAGCGGCACAGCCCCCGCCCCCATG 60-

GCCCGTCGGAGCGCCAGAGGCCAGAGGCGGGAGGGGCGGAGGGCCCGGGGGAGCCCGGCACCCGCTG 168
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
GCCCGTCGGAGCGCCAGAGGCCAGAGGCGGGAGGGGCGGAGGGCCCGGGGGAGCCCGGCACCCGCTG 120

CCGCTCGCGCTGCCTGGGCCTTGGCCTGGCGCTGCCTGGCGCTGCCTGGCGCTGCTGGCCGTG 228
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
CCGCTCGCGCTGCCTGGGCCTTGGCCTGGCGCTGCCTGGCGCTGCCTGGCGCTGCTGGCCGTG 180

AGTTTGGGGAGCCGGGCATCGCTGTCCCGCCCCGGCCTGCCTGCCCAGGAGCTGCCAGGAGAGTGGTG 288
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
AGTTTGGGGAGCCGGGCATCGCTGTCCCGCCCCGGCCTGCCTGCCCAGGAGCTGCCAGGAGAGTGGTG 240

AGAGGAGGACCAGAGCGGACCCCGTCGGAACTGAATCCCCAGAACTCCCAGAGAAAGCCAGGATCCT 348
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
GAGGAGGACCAGGAGCGGACCCCGTCGGAACTGAATCCCCAGAACTCCCAGAGAAAGCCAGGATCCT 300

CCCTTTCCTGAACCGACTAGTTCGGCCTGCAGAAGTGCACCTAAAGGCCGGAAAAACA 408
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCTTTCCTGAACCGACTAGTTCGGCCTGCAGAAGTGCACCTAAAGGCCGGAAAAACA 360

'0.
'0.1   GI:37182857
JA.

apiens (human)
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 1353)
H.F., Gurney,A.L., Abaya,E., Baker,K., Baldwin,D., Brush,J.,
., Chow,B., Chui,C., Crowley,C., Currell,B., Deuel,B.,
, Eaton,D., Foster,J., Grimaldi,C., Gu,Q., Hass,P.E.,
s., Huang,A., Kim,H.S., Klimowski,L., Jin,Y., Johnson,S.,
Lewis,L. Liao.D., Mark,M., Robbie,E., Sanchez,C.,
eld,J., Seshagiri,S., Simmons,L., Singh,J., Smith,V.,
n,J., Vagts,A., Vandlen,R., Watanabe,C., Wieand,D., Woods,K.,
H., Yansura,D., Yi,S., Yu,G., Yuan,J., Zhang,M., Zhang,Z.,
A., Wood,W.I and Godowski,P.
creted Protein Discovery Initiative (SPDI), a Large-Scale
to Identify Novel Human Secreted and Transmembrane Proteins:
nformatics Assessment
Res. 13 (10), 2265-2270 (2003)
09
ses 1 to 1353)
H.F.
Submission
ted (01-AUG-2003) Department of Bioinformatics, Genentech,
1 DNA Way, South San Francisco, CA 94080, USA
Location/Qualifiers
1..1353
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DNA30879"
1..1353
/locus_tag="UNQ181"
58..807
/locus_tag="UNQ181"
/note="PRO207"
/codon_start=1
/product="Apo-3 Ligand"
/protein_id="AAQ89229.1"
/db_xref="GI:3718258"
/translation="MAARRSQRRGRRGEPGTALLVPLALGLGLALACLGLLLAVVSL
GSRASLSAQEPAQEELVAEEDQDPSELNPGTEESQDPAPFLNRLVRPRRSAAPKGRKTR
ARRAIAAHYHVHPRPGQDGAQAGVDVGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGL
YYLYCVHFDEGKAVYLKDLLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLAL
RPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH"

arity   85.4%;   Score 1172;   DB 9;   Length 1353;
onservative 99.8%;   Pred. No. 0;   Mismatches   3;   Indels   0;   Gaps   0;

```
Qy  349 GCGCCTTTCCTGAACCGACTAGTTGGCCTTCGCAGAAGTGCACCTAAAGGCCGAA
            |||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db  301 GCGCCTTTCCTGAACCGACTAGTTCCGCCTTCGCAGAAGTGCACCTAAAGGCCGAA

Qy  409 CGGGCTCGAAGAGGCGGATCGCAGCCCCATTATGAAGTTCATCCACGACCTGGACAGGA
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 CGGGCTCGAAGAGGCGGATCGCAGCCCCATTATGAAGTTCATCCACGACCTGGACAGGA

Qy  469 GCGCAGGCAGGTGTGGACGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAG
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GCGCAGGCAGGTGTGGACGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAG

Qy  529 AGCCCTCGCGCTACAACCGCCAGATCGGGAGTTTATAGTCACCGGGCTGGGCT
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 AGCCCTCGCGCTACAACCGCCAGATCGGGAGTTTATAGTCACCGGGCTGGGCT

Qy  589 TACCTGTACTGTCAGGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGGA
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 TACCTGTACTGTCAGGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGGA

Qy  649 CTGGTGGATGGTGTGCTGGCCCTGGCCTGCCAGGTGTCTGGGCTGTTGGGCCTGGGGGC
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 CTGGTGGATGGTGTGCTGGCCCTGGCCTGCCAGGTGTCTGGGCTGTTGGGCCTGGGGGC

Qy  709 TCCCTGGGCCCCAGTCTCCGCCTCTGCCCAGGTGTCTGGGCCATCTCAAGGCTGCCCCCTTCC
            |||||||| ||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db  661 TCCCTGGGGCCCCAGTCTCCGCCTCCGCCTCTGCCCAGGTGTCTGGGCCATCTCAAGGCTGCCCCCTTCC

Qy  769 TCCTCCCTGCGGATCCGGACACCCCCTGGGCCCTCACTGAGGGGCCCTGTCTCCCCAGTCGTCCCA
            ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 TCCTCCCTGCGGATCCGGACACCCCCTGGGCCCTCAAGGCTGCCCCCTTCCAGTCGTCCCA

Qy  829 TACTTCGGACTCTTCAGGTTCACTGAGGGGCACCCCGGTCTCCCGTCTCCCCGCAGTCGTCCCA
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 TACTTCGGACTCTTCAGGTTCACTGAGGGGCACCCCGGTCTCCCGTCTCCCCGCAGTCGTCCCA

Qy  889 GCCGGCTCCCCTCGCAGAGCTCTCTGGGCACCCGGTCCCCTGCCCACCCTCAGC
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 GCCGGCTCCCCTCGCAGAGCTCTCTGGGCACCCGGTCCCCTGCCCACCCTCAGC

Qy  949 CTTTGCTCCAGACCTGCCCCTCCTCTAGAGGCTGTCACTGTTCACGTGTTT
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 CTTTGCTCCAGACCTGCCCCTCCTCTAGAGGCTGTCACTGTTCACGTGTTT

Qy 1009 TCCCAATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGGCCACTC
            |||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 TCCCACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGGCCACTC

Qy 1069 CCTCACTAGCTCCCCAATCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCCCC
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 CCTCACTAGCTCCCCAATCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCCCC

Qy 1129 GCCACAGACCCCCAGGGCATTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCA
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 GCCACAGACCCCCAGGGCATTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCA

Qy 1189 ACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGCAGGAGAAGCCAAAGAGACT
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 ACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGCAGGAGAAGCCAAAGAGACT

Qy 1249 CTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAAACAAGACTCCTCCCCTT(
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 CTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAAACAAGACTCCTCCCCTT(

Qy 1309 TTCCCTGTGGATTTTAAAACAGATATTATTATTATTGTGACAAAATGT
            |||||||| ||||||| |||||||||||||||   |||||||||||||
Db 1261 TTCCCTGTGTGGATTTTTAAAACAGATATTATTTTTATTATTATTGTGACAAAATGT

Qy 1369 AATGG 1373
            |||||
Db 1321 AATGG 1325
```

```
CCCTCGGGTCCCGGGATGGGGGGGCGTGAGGCGGCAGGCAGCAGCCCCCGGCCCCCCCATG 108
||||| |||||| ||||||||||||||||||||||||||||||||||||||||||||||||
CCCTCCGGGTCCCCGGGATGGGGGGGCGTGAGGCGGCAGGCAGCAGCCCCCGGCCCCCCCATG 60

CCCGTCTGGAGCCAGAGGCGGAGGGGGCCGCCGGGGGGAGCCGGGGAGCACCGCCCTGCTG 168
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCGTCTGGAGCCAGAGGCGGAGGGGGCCGCCGGGGGGAGCCGGGGAGCACCGCCCTGCTG 120

CGCTCGCGCCTGGGCGCTGGGCGCTGGCCGCGGGCCTGCCTGCCTGCCGCCTTCCTGCCTGCTG 228
||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
CGCTCGCGCTGGGCGCTGGGCGCTGGCCGCGGGCCTGCCTGCCTGCCGCCTTCCTGCCTGCTG 180

AGTTTGGGAGCCGGGCATCGCTGCTCCGCCCAGGAGCTGCCCCAGGAGGAGCTGTGTG 288
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGTTTGGGAGCCGGGCATCGCTGCTCCGCCCAGGAGCTGCCCCAGGAGGAGCTGTGTG 240

AGGAGGACCAGGACCCGCTGGAACTGAATCCCCAGACACAGAAGAAAGCCAGGATCCT 348
||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
AGGAGGACCAGGACCCGTCGGAACTGAATCCCCAGACACAGAAGAAAGCCAGGATCCT 300
```

872   1368 bp   mRNA   linear   PRI 04-MAY-1998
sapiens Apo3/DR3 ligand (APO3L) mRNA, complete cds.
872
872.1   GI:3108230

sapiens (human)
sapiens
yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
lia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ases 1 to 1368)
ers,S.A., Sheridan,J.P., Pitti,R.M., Brush,J., Goddard,A. and
nazi,A;
ification of a ligand for the death-domain-containing receptor

  Biol. 8 (9), 525-528 (1998)
355
43
ases 1 to 1368)
ers,S.A., Sheridan,J.P., Pitti,R.M., Brush,J., Goddard,A. and
nazi,A.
t Submission
tted (25-MAR-1998) Molecular Oncology, Genentech, 1 DNA Way,
San Francisco, CA 94080, USA
Location/Qualifiers
1..1368
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /chromosome="17"
  /map="17p13"
1..1368
  /gene="APO3L"
58..807
  /gene="APO3L"
  /function="binds to the death-domain-containing receptor
Apo3/DR3"
  /note="TNF homolog"
  /codon_start=1
  /product="Apo3/DR3 ligand"
  /protein_id="AAC39724.1"
  /db_xref="GI:3108231"
  /translation="MAARSQRRGRRGEPGTALLVPLALGLGLALACIGLLIAVVSL
GSRASLSAQEPAQEELVAEEDQDPSELNPGTEESGDPAPFLNRLVRPRRSAPKGRKTR
ARRATAAHYRVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYNRQIGBFIVTRAGL
YYLYCQVHFDEGKAVYLKLDLLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLAL
RPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH"

larity   85.4%;   Score 1172;   DB 9;   Length 1368;
larity   99.8%;   Pred. No. 0;
Conservative   0;   Mismatches   3;   Indels   0;   Gaps   0;

241 GCAGAGGAGGACCAGGACCCGTCGGAACTGAATCCCCAGACAGAGAAAGCCAGG

Db

349 GCGCCTTCTGAACCGACTAGTTCGGCCTCGCCGAGAGTGCACCTAAAGGCCGGA
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
301 GCGCCTTCTGAACCGACTAGTTCGGCCTCGCCGAGAGTGCACCTAAAGGCCGGA

Qy
Db

409 CGGGCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGACAGG
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
361 CGGGCTCGAAGAGAGGATCGCAGCCCATTATGAAGTTCATCCACGACCTGACAGG

Qy
Db

469 GCGCAGGCAGTGTGACGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACA
     |||||||||||||||||||||||||||||||||||||||||||||||||
421 GCGCAGGCAGTGTGACGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACA

Qy
Db

529 AGCCCCTGGCTACAACCGCCAGATCGGGAGTTTATAGTGCACCGGGCTGGGC
     |||||||||||||||||||||||||||||||||||||||||||||||||||
481 AGCCCCTGGCTACAACCGCCAGATCGGGAGTTTATAGTGCACCGGGCTGGGC

Qy
Db

589 TACCTGTACTGTCAGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGG
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
541 TACCTGTACTGTCAGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGG

Qy
Db

649 CTGGTGGATGGTGTGCTGGCCCTGCCTGCGCTGGTGTCTGGCCTGGCCTGCGG
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
601 CTGGTGGATGGTGTGCTGGCCCTGCCTGCGCTGGTGTCTGGCCTGGCCTGCGG

Qy
Db

709 TCCCTCGGGCCCCAGCTCCGCCTCTGCCAGGTGTCTGGGCTGTTGGCCTGCGGC
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
661 TCCCTCGGGCCCCAGCTCCGCCTCTGCCAGGTGTCTGGGCTGTTGGCCTGCGGC

Qy
Db

769 TCCTCCTGCGATCCGCACCCTCCCCTGCGGCCCATCTCAAGGCTGCCCCCTTCC
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
721 TCCTCCTGCGATCCGCACCCTCCCCTGCGGCCCATCTCAAGGCTGCCCCCTTCC

Qy
Db

829 TACTTCGGACTCTTCCAGGTTCACTGAGGGCCCCTGGTCTCCCCACAGTGTCCC
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
781 TACTTCGGACTCTTCCAGGTTCACTGAGGGCCCCTGGTCTCCCCGCAGTGTCCC

Qy
Db

889 GCCGGCTCCCCTCGACAGCTCTCTGGCACCCGGTCCCCTGCCCCACCCTCAC
     |||||||||||||||||||||||||||||||||||||||||||||||||||
841 GCCGGCTCCCCTCGACAGCTCTCTGGCACCCGGTCCCCTGCCCCACCCTCAC

Qy
Db

949 CTTTGCTCCAGACCTGCCCCCTCCCTCCTCTAGAGGCTGCCTGGGCCTTGTTCACGTGT
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
901 CTTTGCTCCAGACCTGCCCCCTCCCTCCTCTAGAGGCTGCCTGGGCCTTGTTCACGTGT

Qy
Db

1009 TCCCACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCCACCGCCAC
      ||||||||||||||||||||||||||||||||||||||||||||||||||||
961   TCCCACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCCACCGCCAC

Qy
Db

1069 CCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCC
      |||||||||||||||||||||||||||||||||||||||||||||||||||||
1021 CCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCC

Qy
Db

1129 GCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGTCT
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
1081 GCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGTCT

Qy
Db

1189 ACCCCACTTCAGGCACTAAGAGAGGGGCTGGACCTGGCCGGCAGGAAGCCAAAGAGA
      |||||||||||||||||||||||||||||||||||||||||||||||||
1141 ACCCCACTTCAGGCACTAAGAGAGGGGCTGGACCTGGCCGGCAGGAAGCCAAAGAGA

Qy
Db

1249 CTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAAACAAGACAAGCTCCCCT
      |||||||||||||||||||||||||||||||||||||||||||||||||||||
1201 CTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAAACAAGACAAGCTCCCCT

Qy
Db

1309 TTCCCTGTGGATTTTTAAAACAGATATTATTTTATTATTGTGACAAAATG
      ||||||||||||||||||||||||||||||||||||||||||||||||||
1261 TTCCCTGTGGATTTTTAAAACAGATATTATTTTATTATTGTGACAAAATG

Qy
Db

1369 AATGG 1373
1321 AATGG 1325

Qy
Db

```
Db   515 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         AGCCCTCTGCGCTACAACCGCCAGATCGGGAGTTTATAGTCACCCGGCGGCTCGGGCT...

Qy   589 TACCTGTACTGTCAGGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAGCTGGA...
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   575 TACCTGTACTGTCAGGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAGCTGGA...

Qy   649 CTGGTGGATGGTGTGCTGGCCCTGCCTGCCTGAGGAGAATTCTCAGCCACTGCGGCC...
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   635 CTGGTGGATGGTGTGCTGGCCCTGCCTGCCTGAGGAGAATTCTCAGCCACTGCGGCC...

Qy   709 TCCCCTGGGGCCCAGTCCTCGCCCTCGCCAGTGTCTCGGGCTGTTGGCCCTGGGGCCC...
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   695 TCCCCTGGGGCCCAGTCCTCGCCCTCGCCAGTGTCTCGGGCTGTTGGCCCTGGGGCCC...

Qy   769 TCCTCCCTGCGGATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCT...
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   755 TCCTCCCTGCGGATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCT...

Qy   829 TACTTCGGACTCTTCCAGGTTCACTGAGGGCCCTGGTCTCCCCACAGTCGTCCCA...
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   815 TACTTCGGACTCTTCCAGGTTCACTGAGGGCCCTGGTCTCCCCGCAGTCGTCCCA...

Qy   889 GCCGGCTCCCCTGACAGCTTCTCTGGGCACCGGTCCCCTGCCCACCCTCAGC...
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   875 GCCGGCTCCCCTGACAGCTTCTCTGGGCACCGGTCCCCTGCCCACCCTCAGC...

Qy   949 CTTTGCTCCAGACCTGCCCCTCCCTCCTAGAGGGCTGCCTGGGCCTGTTCAGTGTTT...
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   935 CTTTGCTCCAGACCTGCCCCTCCCTCCTAGAGGGCTGCCTGGGCCTGTTCAGTGTTT...

Qy  1009 TCCCACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCCACTC...
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   995 TCCCACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCCACTC...

Qy  1069 CCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCCCC...
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1055 CCTCACTAGCTCCCCAATCCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCCCC...

Qy  1129 GCCACAGACCCCCAGGCATTGTGTTCACTGTACTTCGTGTGGGCAAGGATGGGTCCA...
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1115 GCCACAGACCCCCAGTCATTGTGTTCACTGTACTTCGTGTGGGCAAGGATGGGTCCA...

Qy  1189 ACCCCACTTCAGGCACTAAGAGGGCTGGACCTGGGCGCGGCAGGAAGCCAAGAGACT...
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1175 ACCCCACTTCAGGCACTAAGAGGGCTGGACCTGGGCGCGGCAGGAAGCCAAGAGACT...

Qy  1249 CTAGGCCAGGAGTTCCAAATGTGAGGGGCCAGAAACAAGACAAGCTCCTCCCTTG...
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1235 CTAGGCCAGGAGTTCCAAATGTGAGGGGCCAGAAACAAGACAAGCTCCTCCCTTG...

Qy  1309 TTCCCTGTGATTTTTAAAACAGATATTATTTTATTATTTTATTGTGACAAAATGTT...
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1295 TTCCCTGTGATTTTTAAAACAGATATTATTTTATTATTTTATTGTGACAAAATGTT...

Qy  1369 AATGG  1373
         |||||
Db  1355 AATGG  1359
```

e
: US 6207642-A 1 27-MAR-2001;
Location/Qualifiers
1..1236
/organism="unknown"
/mol_type="unassigned DNA"

```
                    69.8%;   Score 958;   DB 6;  Length 1236;
arity      99.6%;   Pred. No. 0;
onservative    0;   Mismatches    5;   Indels    0;   Gaps    0;
```

```
Qy   129  GCGGAGGGGGCGCGCGGGGGGGAGCCGGGCACCGCCCCTGCTCGTCCGCTCGGCGCTG   188
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    24  GCGGAGGGGGGCGCGGGGGGGAGCCGGGCACCGCCCCTGCTCGTCCCCGCTCGGCGCTG    83
```

(sequence alignment block)

RESULT 8
BD057124
LOCUS   BD057124   1236 bp   DNA   linear   PAT 2'
DEFINITION Member of the tnf family useful for treatment and diagnos
            disease.
ACCESSION   BD057124
VERSION   BD057124.1 GI:22602730
KEYWORDS   JP 2001513626-A/1.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutel
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
REFERENCE   1 (bases 1 to 1236)
  AUTHORS   Wiley,S.R.
  TITLE   Member of the tnf family useful for treatment and diagnos
  JOURNAL   Patent: JP 2001513626-A 1 04-SEP-2001;
            ABBOTT LABORATORIES
COMMENT   PN   JP 2001513626-A/1
            PD   04-SEP-2001
            PF   12-FEB-1998 JP 1998535077
            PR   12-FEB-1997 US 08/798692,10-FEB-1998 US   09/021
            STEVEN R WILEY
            PC   C12Q1/68,C12N15/63,C07K14/525,C07K16//28,G01N33/50,G0
            PC   A61K38/19
            CC   Strandedness: Single;
            CC   Topology: Linear;
            FH   Key   Location/Qualifiers.
FEATURES   Location/Qualifiers
     source   1..1236
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match   69.8%;   Score 958;   DB 6; Length 1236;
Best Local Similarity 99.6%;   Pred. No. 0;
Matches 1208;   Conservative   0;   Mismatches   5;   Indels   0;
```
Qy   129  GCGGAGGGGGCGCGCGGGGGGGAGCCGGGCACCGCCCCTGCTCGTCCGCTCGGCGCTG   188
Db    24  GCGGAGGGGGGCGCGGGGGGGAGCCGGGCACCGCCCCTGCTCGTCCCCGCTCGGCGCTG    83
Qy   189  GGGCCTGGCGCTGGCGCCTGGCGCCTGGCGCCTGGCGCCTGGCGCTGTGCGGGGAGC   248
Db    84  GGGCCTGGCGCTGGCGCCTGGCGCCTGGCGCCTGGCGCGTGGCGGGGAGC   143
Qy   249  ATCGCTGTCGCCGCCCAGGAGCCTGCCCAGGAGCCTGGTGCGCAGAGGAGGACCAC   308
Db   144  ATCGCTGTCGCCCGCCCAGGAGCCTGCCCAGGAGCCTGGTGCGCAGAGAGGAGGACCAG   203
```

RESULT 9
BC019047

LOCUS   BC019047   1642 bp   mRNA   linear   PRI 04
DEFINITION Homo sapiens tumor necrosis factor (ligand) superfamily, m
            transcript variant 2, mRNA (cDNA clone MGC:20669 IMAGE:476
            complete cds.
ACCESSION   BC019047
VERSION   BC019047.2 GI:33874709
KEYWORDS   MGC.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1642)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M.,   Schu
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bha
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsie
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Mye
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE   Generation and initial analysis of more than 15,000 full-l
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2 (bases 1 to 1642)
  AUTHORS   Strausberg,R.
  TITLE   Direct Submission
  JOURNAL   Submitted (07-DEC-2001) National Institutes of Health, Man
            Gene Collection (MGC), Cancer Genomics Office, National Ca
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892
            USA
REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   On Aug 19, 2003 this sequence version replaced gi:1751213(
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis Staudt
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramura
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Broc
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L, Karlins,E., Kwong,P., Laric,P., Legas
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McClo
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touch
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggin
            Young,A., Zhang,L.-H. and Green,E.D.
            Clone distribution: MGC clone distribution information car
            through the I.M.A.G.E. Consortium/LLNL at: http://image.l
            Series: IRAL Plate: 30 Row: p Column: 5
            This clone was selected for full length sequencing because
            passed the following selection criteria: matched mRNA gi:
FEATURES   Location/Qualifiers
     source   1..1642

BAACTGAATCCCCAGACAGAGAAGAAAGCCAGATCTCTGCGCCTTTCCTGACCGACT 368
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
BAACTGAATCCCCAGACAGAGAAGAAAGCCAGATCTCTGCGCCTTTCCTGACCGACT 263

CGGCCTCGCGAAGTGCACCTAAAGGCCGGAAAACACGGGCTCGAAGAGGCGATCGC 428
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CGGCCTCGCGAAGTGCACCTAAAGGCCGGAAAACACGGGCTCGAAGAGGCGATCGC 323

CATTATGAAGTTCATCCACGACCTGACAGGACGGAGCGCGCAGGCAGTGTCGACGG 488
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CATTATGAAGTTCATCCACGACCTGACAGGACGGAGCGCGCAGGCAGTGTCGACGG 383

GTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAGCCCTCTGCGCTACAACCG 548
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAGCCCTCTGCGCTACAACCG 443

ATCGGGGAGTTTATAGTCACCCGGGCGTGGGCTCTACTACCTGTACTGTCAGGTGCA 608
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTACTACCTGTACTGTCAGGTGCA 503

GATGAGGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGCATGGTGTGCTGGC 668
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GATGAGGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGCATGGTGTGCTGGC 563

CGCTGCCTCGGGAGGAATTCTCAGCCACTGCGGCCAGTTCCCTGGGGCCCCAGCTCCG 728
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CGCTGCCTCGGGAGGAATTCTCAGCCACTGCGGCCAGTTCCCTGGGGCCCCCAGCTCCG 623

TGCCAGGTGTCTGGGCTGTTGGGCCCTGCGCTGCGGGTGCTCCCTGCGGATCCGCAC 788
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGCCAGGTGTCTGGGCTGTTGGGCCCTGCGCTGCGGGTGCTCCCTGCGGATCCGCAC 683

CCCTGGGCCCATCTCAAGCTGCGCCCCCTTCCTCACCTACTTCGGACTCTTCCAGGT 848
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCTGGGCCCATCTCAAGCTGCGCCCCCTTCCTCACCTACTTCGGACTCTTCCAGGT 743

TGAGGGGCCCTGTCTCCCCACAGTCGTCCCGCAGTCGTCCCGGCTGCCGGCTCCCCTGACAGCT 908
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGAGGGGCCCTGTCTCCCCACAGTCGTCCCGCAGTCGTCCCGGCTGCCGGCTCCCCTGACAGCT 803

GGGCACACCGGTCCCCCTGCGCCCACCACCCTGCGCCTCTTTGCTCCAGACCGTGCCCCC 968
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGGCACACCGGTCCCCCTGCGCCCACCACCCTGCGCCTCTTTGCTCCAGACCGTGCCCCC 863

TCTAGAGGCTGCCTGGGCCTGTTCACGTGTTTTCATCCACATAAATACAGTATT 1028
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCTAGAGGCTGCCTGGGCCTGTTCACGTGTTTTCATCCACATAAATACAGTATT 923

CTCTTATCTTACAACTCCCCACCGCCCACTCTTCCACCTCACTAGCTCCCCAATCC 1088
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTCTTATCTTACACAACCCCACCGCCCACTCTTCCACCTCACTAGCTCCCCAATCC 983

CCCTTTGAGGGCCCCCAGTGATCCGACTCCCCCCTGGCCACAGACCCCCAGGGCAT 1148
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCTTTGAGGGCCCCCAGTGATCTGACTCCCCCCTGGCCACAGACCCCCAGGGCAT 1043

TTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCCCACTTCAGGCACTAAG 1208
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
TTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCCCACTTCAGGCACTAAG 1103

GCTGGACCTGGCGGCGAGGAAGCCAAAGAGACTGGCCTAGGCCAGGAGTTCCCAAA 1268
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCTGGACCTGGCGGCGAGGAAGCCAAAGAGACTGGCCTAGGCCAGGAGTTCCCAAA 1163

GAGGGGCGGAGAAACAAGACAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTAAAA 1328
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGGGGGCGAGAAACAAGACAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTTAAAA 1223

ATTATTATTT 1341
       ||||||||||
ATTATTATTT 1236

```
Qy   1136 ACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGTCTCCAGAAGACC
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1366 ACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGTCTCCAGAAGACC

Qy   1196 TTCAGGCACTAAGAGAGGGCTGACCTGGCGGCGGAGGAGACCAAAGAGACTGGGCCTA
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1426 TTCAGGCACTAAGAGAGGGCTGACCTGGCGGCGGAGGAGACCAAAGAGACTGGGCCTA

Qy   1256 AGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCCTTGAGAATTC
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1486 AGGAGTTCCCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCCTTGAGAATTC

Qy   1316 TGGATTTTTAAAACAGATATTATTTTTATTATTATTGTGACAAAATGTTGATAAAT
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1546 TGGATTTTTAAAACAGATATTATTTTTATTATTATTGTGACAAAATGTTGATAAAT
```

RESULT 10
AX180714
LOCUS   AX180714   898 bp   DNA   linear   PAT 06
DEFINITION   Sequence 1 from Patent WO0145730.
ACCESSION   AX180714
VERSION   AX180714.1   GI:15132570
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM   synthetic construct
             artificial sequences.
REFERENCE   1
  AUTHORS   Wiley,S.R.
  TITLE   Tweak receptor
  JOURNAL   Patent: WO 0145730-A 1 28-JUN-2001;
             IMMUNEX CORPORATION (US)
FEATURES   Location/Qualifiers
     source   1..898
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
     CDS   52..873
                     /note="unnamed protein product; human TWEAK fusic
                     construct"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC50004.1"
                     /db_xref="GI:15132571"
                     /db_xref="REMTREMBL:CAC50004"
                     /translation="MATGSRTSLLLAFGLLCLPWLQEGSATSSDRMKQ
                     KIYIHENEIARIKKLIGERTRSSLGSRASLSAQEPAQEELVAEEDQDP
                     QDPAPFLNRLVRPRSAPKGRKTRARRAIAAHYEVHPRPGQDGAQAGVI
                     RINSSSPLRYNRQIGEPIVTRAGLYYLYCQVHFDEGKAVYLKLDLLVDV
                     FSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFLTY"

ORIGIN

Query Match   45.5%;   Score 625;   DB 6;   Length 898;
Best Local Similarity 100.0%;   Pred. No. 0;
Matches   625;   Conservative   0;   Mismatches   0;   Indels   0;   G

```
Qy   232 AGTTTGGGGAGCCGGGCATCGCTGTCCGCCCAGGAGAGCTGCCCAGGAGGAGCTGG
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   250 AGTTTGGGGAGCCGGGCATCGCTGTCCGCCCAGGAGAGCTGCCCAGGAGGAGCTGG

Qy   292 GAGGAGGACCAGGACCCGTCGGAACTGAATCCCCAGACAGAGGAAAAGCCAGGATC
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   310 GAGGAGGACCAGGACCCGTCGGAACTGAATCCCCAGACAGAGGAAAAGCCAGGATC

Qy   352 CCTTTCCTGAACCGACTAGTTCGGCCTCGCCTCGCAGAGTGCACCTAAAGGCCGAAAA
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   370 CCTTTCCTGAACCGACTAGTTCGGCCTCGCCTCGCAGAGTGCACCTAAAGGCCGAAAA

Qy   412 GCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACG
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   430 GCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACG

Qy   472 CAGGCAGGTGTGGACGGGACAGTGAGTGGCTGGGAGGAGCCAGAATCAACAGCT
```

arity   -58.0%;   Score 796;   DB 9;   Length 1642;   Pred. No. 0;
onservative   0;   Mismatches   2;   Indels   0;   Gaps   0;

```
Qy   TGTGGACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAGCCCTC  535
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   TGTGGACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAGCCCTC  765

Qy   CTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTACTACCTGT  595
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   CTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTACTACCTGT  825

Qy   TCAGGTGCACTTTGATGAGGAGGCTGCTCTACCTGAAGCTGACTGCTGGTGG  655
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   TCAGGTGCACTTTGATGAGGAGGCTGCTCTACCTGAAGCTGACTGCTGGTGG  885

Qy   TGTGCTGGCCCTGGCGCTGCTGGAGGAATTCTCAGCCACTGCGGCCAGTTCCCTCG  715
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   TGTGCTGGCCCTGGCGCTGCTGGAGGAATTCTCAGCCACTGCGGCCAGTTCCCTCG  945

Qy   CCAGCTCCGCCTCTGCCAGGTGTCTCGGCTGTTGGCCCTGCGGCCAGGGTCCTCCC  775
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   CCAGCTCCGCCTCTGCCAGGTGTCTCGGCTGTTGGCCCTGCGGCCAGGGTCCTCCC  1005

Qy   GATCCGCACCCCCTGGGCCCATCTCAAGGCGCCCCCTTCCTCACCTACTTCG  835
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   GATCCGCACCCCCTGGGCCCATCTCAAGGCGCCCCCTTCCTCACCTACTTCG  1065

Qy   CCTTCCAGGTTCACTGAGGGGCCCTGGCGCCCGGCTCTCCCCAGTCGTCCCCAGGCTGCCGGCT  895
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   CCTTCCAGGTTCACTGAGGGGCCCTGGCGCCCGGCTCTCCCCAGTCGTCCCCAGGCTGCCGGCT  1125

Qy   TCGACAGCTCTGGGCACCCGGTCCCCTCTGCCCCACCCTCAGCCGCTCTTTGCT  955
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   TCGACAGCTCTGGGCACCCGGTCCCCTCTGCCCCACCCTCAGCCGCTCTTTGCT  1185

Qy   GACCTGCCCCCCCCTCTAGAGGCTGCCTGGGCCTGGCCTGTTCACGTGTTTTCATCCCACA  1015
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   GACCTGCCCCCCCCTCTAGAGGCTGCCTGGGCCTGGCCTGTTCACGTGTTTTCATCCCACA  1245

Qy   ATACAGTATTCCCACTCTTATCTTACAACTCCCCCACCGCCCACTCTCCACCTCACT  1075
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   ATACAGTATTCCCACTCTTATCTTACAACTCCCCCACCGCCCACTCTCCACCTCACT  1305

Qy   TCCCCAATCCCTGACCCTTTGAGGCCCCCAGTGATCTGACTCGACTCCCCCCTGCCACAG  1135
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   TCCCCAATCCCTGACCCTTTGAGGCCCCCAGTGATCTGACTCGACTCCCCCCTGCCACAG  1365

Qy   CAGGCAGGTGTGGACGGGACAGTGAGTGGCTGGGAGGAGCCAGAATCAACAGCT
```

```
|||||||||||||||||||||||||||||||||||||||||||||||||||
AGGTGTGGACGGGACAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTCCAGC  549
|||||||||||||||||||||||||||||||||||||||||||||||||||
3CGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCGGGCTGGGCTCTACTAC  591
|||||||||||||||||||||||||||||||||||||||||||||||||||
3CGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCGGGCTGGGCTCTACTAC  609
|||||||||||||||||||||||||||||||||||||||||||||||||||
CTGTCAGGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGACTTGCTG  651
|||||||||||||||||||||||||||||||||||||||||||||||||||
CTGTCAGGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGACTTGCTG  669
|||||||||||||||||||||||||||||||||||||||||||||||||||
TGGTGTGCTGGCCCTGGCCTGCTGCCTGGAGGAATTCTCAGCCACTGCGCCCAGTTCC  711
|||||||||||||||||||||||||||||||||||||||||||||||||||
TGGTGTGCTGGCCCTGGCCTGCTGCCTGGAGGAATTCTCAGCCACTGCGCCCAGTTCC  729
|||||||||||||||||||||||||||||||||||||||||||||||||||
GCCCCAGCTCCGCCTCTGCCAGTGTCTGGGCGTTGCCCTGCGCGGCCAGGGTCC  771
|||||||||||||||||||||||||||||||||||||||||||||||||||
GCCCCAGCTCCGCCTCTGCCAGTGTCTGGGCGTTGCCCTGCGCGGCCAGGGTCC  789
|||||||||||||||||||||||||||||||||||||||||||||||||||
GCGGATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACCTAC  831
|||||||||||||||||||||||||||||||||||||||||||||||||||
GCGGATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACCTAC  849
|||||||||||||||||||||||||||||||||||||||||||||||||||
ACTCTTCCAGGTTCACTGAG  856
||||||||||||||||||||
ACTCTTCCAGGTTCACTGAG  874
```

```
6                 60268 bp    DNA    linear    PRI 31-OCT-2002
piens chromosome 17, clone RP11-186B7, complete sequence.
6
6.10  GI:24431829

piens (human)
piens
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 60268)
B., Nusbaum,C. and Lander,E.
piens chromosome 17, clone RP11-186B7
shed
es 1 to 60268)
B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B.,
., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
a,P., FitzHugh,W., Forrest,I., Funke,R., Gage,D., Horton,L.,
,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Y,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
n,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu.X.,
., Ye,W.J., Zimmer,A. and Zody,M.
Submission
ed (08-DEC-1999) Whitehead Institute/MIT Center for Genome
h, 320 Charles Street, Cambridge, MA 02141, USA
es 1 to 60268)
B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
I., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
a,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
MacLean,C., Macdonald,P., Major,J., Matthews,C.,
y,M., Meldrim,J., Meneus,L., Mihova,V., Mlenga,V.,
T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
```

O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rog
Roman,J., Roy,A., Schauer,S., Schupback.R., Seaman,S., Seve
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Tal
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H
Viel.R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainc
Zembek,L., Zimmer,A. and Zody,M.

TITLE   Direct Submission
JOURNAL Submitted (08-OCT-2002) Whitehead Institute/MIT Center for
         Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 4 (bases 1 to 60268)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Andersc
         Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter
         Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.
         Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodg
         Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
         Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
         Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,
         Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
         Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.
         Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
         Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu
         Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,C
         Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
         Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,
         Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,
         Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,
         Travers,M., Vassiliev,H., Viel.R., Vo,A., Wilson,B., Wu,X.,
         Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zc
         Direct Submission
TITLE   Submitted (31-OCT-2002) Whitehead Institute/MIT Center for
JOURNAL Research, 320 Charles Street, Cambridge, MA 02141 USA
COMMENT On Oct 31, 2002 this sequence version replaced gi:23592141.
         All repeats were identified using RepeatMasker:
         Smit, A.F.A. & Green, P. (1996-1997)
         http://ftp.genome.washington.edu/RM/RepeatMasker.html
         --------------- Genome Center
         Center: Whitehead Institute/ MIT Center for Genome Rese
         Center code: WIBR
         Web site: http://www-seq.wi.mit.edu
         Contact: sequence_submissions@genome.wi.mit.edu
         ------------- Project Information
         Center project name: L3849
         Center clone name: 186_B_7

         Only the first 60.3 kilobases of this clone are being subm:
         The remainder overlaps accession number AC113189 [WICGR pro
         L23113].

FEATURES   Location/Qualifiers
     source   1..60268
                  /organism="Homo sapiens"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:9606"
                  /chromosome="17"
                  /map="17"
                  /clone="RP11-186B7"
                  /clone_lib="RPCI-11 Human Male BAC"
     repeat_region complement(7920..8209)
                  /rpt_family="AluSq"
     repeat_region complement(8822..9119)
                  /rpt_family="AluSx"
     repeat_region complement(9356..9663)
                  /rpt_family="AluSq"
     repeat_region complement(12421..12733)
                  /rpt_family="AluSx"
     repeat_region 14360..14438
                  /rpt_family="(TGG)n"
     repeat_region complement(17883..18086)
                  /rpt_family="MIR"
     repeat_region complement(18188..18483)
                  /rpt_family="AluSq"
     repeat_region 22454..22535

repeat_region 41746..42046 /rpt_family="AluSx"
repeat_region 42047..42288 /rpt_family="L2"
repeat_region 42289..42581 /rpt_family="AluSx"
repeat_region 42582..42730 /rpt_family="L2"
repeat_region 43730..43859 /rpt_family="MER94"
repeat_region 44018..44132 /rpt_family="MER21B"

Query Match 45.2%; Score 620; DB 9; Length 60268;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 770; Conservative 0; Mismatches 3; Indels 0; G

```
Qy  601 CAGGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGACTTGCTGGTGGA
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 37278 CAGGTGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGACTTGCTGGTGGA

Qy  661 GTGCTGGCCCTGCGCTGCCTGGAGGAATTCTCAGCGCACTGCGGCCAGTTCCCTCGG
        ||||||||||||||||||||||||||||||||||| ||||||||  ||||||||||
Db 37218 GTGCTGGCCCTGCGCTGCCTGGAGGAATTCTCAGCCACTGCGGCGGAGTTCCCTCGG

Qy  721 CAGCTCCGGCCTCTGCCAGTGTGTCTGGGCTGTTGGCCTGCGGGCCAGGGTCCTCCCT
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 37158 CAGCTCCGGCCTCTGCCAGTGTCTGGTCTGTTGGCCTGCGGGCCAGGGTCCTCCCT

Qy  781 ATCCGCACCCTCCCCTGGGCCATCTCAAGGCTGCCCCCTTCCTCACCTACTTCGC
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 37098 ATCCGCACCCTCCCCTGGGCCATCTCAAGGCTGCCCCTTCCTCACCTACTTCGC

Qy  841 TTCCAGTTCACTGAGGGGCCCTGGTCTCCCCACAGTCGTCCCAGGCTGCCGGGCT
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 37038 TTCCAGTTCACTGAGGGGCCCTGGTCTCCCCGAGTCGTCGTCCCAGGCTGCCGGGCT

Qy  901 CGACAGCTCTTCTGGGCACCCGGTCCCCGTCCCCCTGCCCCACCCTCAGGCGCTCTTTGCT
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 36978 CGACAGCTCTTCTGGGCACCCGGTCCCCGTCCCCCTGCCCCACCCTCAGCCGCTCTTTGCT

Qy  961 CCTGCCCCTCCCTAGAGGCTGCCTGGGCTTCGTGGGCCTGTTTCACGTGTTTTCCATCCCACA
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 36918 CCTGCCCCTCCCTAGAGGCTGCCTGGGCCTGGGCCTGTTTCACGTGTTTTCCATCCCACA

Qy 1021 ACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCCACTCTCCACCTCACT
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 36858 ACAGTATTCCCACTCTTATCTTACAACTCCCCACCGCCCACTCTCCACCTCACT

Qy 1081 CCCAATCCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCCCCCTGGCCACAG
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 36798 CCCAATCCCTGACCCTTTGAGGCCCCCAGTGATCTCGACTCCCCCTGGCCACAG

Qy 1141 CAGGGCATTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCCCAC
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 36738 CAGGGCATTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCCCAC

Qy 1201 GCACTAAGAGAGGGCTGGACCTGGCGGCGGCAAAGGAGACTGGGCCTAGGCC
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 36678 GCACTAAGAGAGGGCTGGACCTGGCGGCGGCAAAGAGACTGGGCCTAGGCC

Qy 1261 TTCCCAAATGTGAGGGCGAGAAACAAGCAGAAGCTCTCCCTTGAGAATTCCCTG
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 36618 TTCCCAAATGTGAGGGCGAGAAACAAGACAAGAAGCTCTCCCTTGAGAATTCCTG

Qy 1321 TTTTAAAACAGATATTATTTTTATTATTATTGTGACAAAATGTTGATAAATGG 1
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 36558 TTTTAAAACAGATATTATTTTTATTATTATTGTGACAAAATGTTGATAAATGG 3
```

RESULT 12
AY081051
LOCUS AY081051 1816 bp mRNA linear PRI
DEFINITION Homo sapiens TWE-PRIL mRNA, complete cds.

/rpt_family="(TA)n"
24279..24530
/rpt_family="MIR"
complement(24517..24655)
/rpt_family="MIR"
24839..24954
/rpt_family="L2"
complement(24973..25268)
/rpt_family="AluY"
complement(25273..25768)
/rpt_family="MER21B"
complement(25769..26057)
/rpt_family="AluSg"
complement(26058..26334)
/rpt_family="MER21B"
26344..26838
/rpt_family="L2"
complement(26839..27139)
/rpt_family="AluSq"
27140..27185
/rpt_family="L2"
27405..27430
/note="<30 qual SNGL region"
27529..27550
/note="single clone coverage"
complement(27552..27689)
/rpt_family="FLAM_C"
27568..27572
/note="<30 qual SNGL region"
complement(27690..28000)
/rpt_family="AluSx"
complement(28005..28312)
/rpt_family="AluSp"
30022..30088
/rpt_family="L2"
30817..30846
/rpt_family="GC_rich"
31023..31072
/rpt_family="(TCC)n"
32341..32378
/rpt_family="(CCG)n"
32571..32601
/rpt_family="GC_rich"
37977..38092
/rpt_family="L1MB5"
38094..38400
/rpt_family="AluSx"
38910..39189
/rpt_family="L2"
complement(39218..39381)
/rpt_family="AluSg/x"
39394..39457
/rpt_family="L2"
39458..39748
/rpt_family="L2"
39593..39597
/note="<30 qual SNGL region"
39749..39759
/rpt_family="L2"
39760..40068
/rpt_family="AluY"
40069..40152
/rpt_family="L2"
40350..40630
/rpt_family="L2"
40718..40905
/rpt_family="MIR"
40926..41025
/rpt_family="L2"
41026..41341
/rpt_family="AluSp"
41345..41745
/rpt_family="L2"

```
Db  361  GGCTCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACGG
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  471  GCAGGCAGTGTGGACGGGACAGCAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTC
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  GCAGGCAGGCAGTGTGGACGGGACAGCAGTGAGTGGCTGGGAGGAAGCCAGAATCAACAGCTC

Qy  531  CCCTCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTA
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  CCCTCTGCGCTACACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTA

Qy  591  CCTGTACTGTCAG  603
              |||||||||||||
Db  541  CCTGTACTGTCAG  553
```

RESULT 13
AC127470
LOCUS   AC127470   218485 bp   DNA   linear   HTG 05-
DEFINITION   Pan troglodytes clone RP43-145D13, WORKING DRAFT SEQUENCE,
             ordered pieces.
ACCESSION   AC127470
VERSION   AC127470.4   GI:31415893
KEYWORDS   HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE   Pan troglodytes (chimpanzee)
  ORGANISM   Pan troglodytes
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1 (bases 1 to 218485)
  AUTHORS   Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjami
             Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Car
             Chu,G., Coleman,B., Coleman,H., Engle,J., Granite,S., Guan
             Gupta,J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
             Hurle,B., Idol,J.R., Karlins,E., Kwong,P., Laric,P., Lee-L
             Legaspi,R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Masi
             Maskeri,B., McDowell,J., Paguirigan,C., Pearson,R., Portno
             Prasad,A., Reddix-Dugue,N., Schandler,K., Schueler,M.G., S
             Sison,C., Stantripop,S., Thomas,J.W., Thomas,P.J., Tsipour
             Vogt,J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E
             NISC Comparative Sequencing Initiative
  TITLE   Unpublished
  JOURNAL   2 (bases 1 to 218485)
REFERENCE   Green,E.D.
  AUTHORS   Direct Submission
  TITLE   Submitted (17-JUL-2002) NIH Intramural Sequencing Center,
  JOURNAL   Grovemont Circle, Gaithersburg, MD 20877, USA
             3 (bases 1 to 218485)
REFERENCE   Green,E.D.
  AUTHORS   Direct Submission
  TITLE   Submitted (05-JUN-2003) NIH Intramural Sequencing Center,
  JOURNAL   Grovemont Circle, Gaithersburg, MD 20877, USA
             On Jun 5, 2003 this sequence version replaced gi:26449071.
COMMENT   ---------- Genome Center
                 Center: NIH Intramural Sequencing Center
                 Center code: NISC
                 Web site: http://www.nisc.nih.gov
                 Contact: nisc_zoo@nhgri.nih.gov
             ---------- Project Information
                 Center project name: cms
                 Center clone name: 145D13

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicat
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8X average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associat
with a Phrap-derived quality score.
             ---------- Summary Statistics

1
1.1 GI:24934973

piens (human)
piens
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 1816)
Balade,B., Medema,J.P., Lopez-Fraga,M., Lozano,J.C.,
oten,G.M., Picard,A., Martinez-A,C., Garcia-Sanz,J.A. and

genous hybrid mRNA encodes TWE-PRIL, a functional cell
TWEAK-APRIL fusion protein
21 (21), 5711-5720 (2002)
9
es 1 to 1816)
Balade,B., Garcia-Sanz,J.A. and Hahne,M.
Submission
ed (21-FEB-2002) DIO, CNB, ctra de Colmenar Viejo, MADRID,
28049, Spain
Location/Qualifiers
      1..1816
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /chromosome="17"
      /note="expressed in primary T lymphocytes and monocytes"
56..1048
      /note="fusion protein; TWEAK/APRIL hybrid; TWEAK and APRIL
      are members of the TNF ligand family, corresponding to
      TNFSF12 and TNFSF13, respectively"
      /codon_start=1
      /product="TWE-PRIL"
      /protein_id="AAL90443.1"
      /db_xref="GI:24934974"
      /translation="MAARRSQRRRGRRGREPGTALLVPLAIGLGLALACLGLLLAVVSL
      GSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTR
      ARRALAAHYEVHPRPGDGAQAGVDGTVSGWEARINSSPLRYNRQIGEFIVTRAGL
      YYLYCQSSDALEAWENGERSRKRRAVLTQKQKKQHSVLHLVPINATSKDDSDVTEVMW
      QPALRRGRGLQAQGYGVRIQDAGVYLLYSQVLFQDVTFTMGQVVSREGQGRQETLFRC
      IRSMPSHPDRAYNSCYSAGVFHLHQGDILSVIIPRARAKLNLSPHGTFLGFVKL"
```

```
arity     40.3%;   Score 553;   DB 9;   Length 1816;
onservative  100.0%;   Pred. No. 8.3e-297;   Indels   0;   Gaps   0;
                 0;   Mismatches   0;

CTCGGGGTCCCGGGATGGGGGGGCGGCGGTGAGGCGGCAGGCACAGCCCCGCCCCCATGGC  110
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTCGGGGTCCCGGGATGGGGGGGCGGCGGTGAGGCGGCAGGCACAGCCCCGCCCCCATGGC   60

CGTCGGAGCCAGAGGCCAGAGGCGGGGGGGGCGCCGAGGGGGGGGCACCGCCCCTGCTGGT  170
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CGTCGGAGCCAGAGGCCAGAGGCGGGGGGGGCGCCGAGGGGGGGGCACCGCCCCTGCTGGT  120

CTCGCGCTGGGCCTGGGCCTGGCGCGCTGGCGCGCTGGCTGCCTCGCCTCCTGCTGCCTGGT  230
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTCGGGCCTGGGCCTGGGCCTGGCGCGCTGGCGCGCTGGCTGCCTCGCCTCCTGCTGCCTGGT  180

TTGGGGAGCCGGGCATCGCTGTCCGGCCCGTCGCCGGCAGGCAGGCAGGAGGAGCCTGGCGGC  290
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TTGGGGAGCCGGGCATCGCTGTCCGGCCCGTCGCCGGCAGGCAGGCAGGAGGAGCCTGGTGGC  240

GAGGACCAGGACCCGTCGGAACTGAATCCCCAGACAGAAGAAAGCCAGGAGATCCTGC  350
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GAGGACCAGGACCCGTCGGAACTGAATCCCCAGACAGAAGAAAGCCAGGAGATCCTGC  300

CTTCCTGAACCGACTAGTTCGGCCTTCGCAGAAGTGCACCTAAAGGCCGAAAACACG  410
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
TTCCTGAACCGACTAGTTCGGCCTTCGCAGAAGTGCACCTAAAGGCCGAAAACACG  360

TCGAAGAGCGATCGCAGCCCATTATGAAGTTCATCCACGACCTGGACAGGACGGAGC  470
||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

equencing vector: plasmid; n/a; 100% of reads
hemistry: Dye-terminator Big Dye; version 0.990319
ssembly program: Phrap; version 0.990319
onsensus quality: 214085 bases at least Q40
onsensus quality: 215255 bases at least Q30
onsensus quality: 216264 bases at least Q20
nsert size: 190000; agarose-fp
nsert size: 216885; sum-of-contigs
uality coverage: 12.65x in Q20 bases; agarose-fp
uality coverage: 11.08x in Q20 bases; sum-of-contigs

:. This is a 'working draft' sequence. It currently
sists of 17 contigs. Gaps between the contigs
represented as runs of N. The order of the pieces
elieved to be correct as given, however the sizes
he gaps between them are based on estimates that have
ided by the submittor.
s sequence will be replaced
he finished sequence as soon as it is available and
accession number will be preserved.

```
1       7548: contig of 7548 bp in length
7549    7648: gap of unknown length
7649   31917: contig of 24269 bp in length
31918  32017: gap of unknown length
32018  50433: contig of 18416 bp in length
50434  50533: gap of unknown length
50534  95274: contig of 44741 bp in length
95275  95374: gap of unknown length
95375  99988: contig of 4614 bp in length
99989 100088: gap of unknown length
00089 109873: contig of 9785 bp in length
09874 109973: gap of unknown length
09974 117619: contig of 7646 bp in length
17620 117719: gap of unknown length
17720 128625: contig of 10906 bp in length
28626 128725: gap of unknown length
28726 143521: contig of 14796 bp in length
43522 143621: gap of unknown length
43622 150832: contig of 7211 bp in length
50833 150932: gap of unknown length
50933 156073: contig of 5141 bp in length
56074 156173: gap of unknown length
56174 187450: contig of 31277 bp in length
87451 187550: gap of unknown length
87551 191471: contig of 3921 bp in length
91472 191571: gap of unknown length
91572 207263: contig of 15692 bp in length
07264 207363: gap of unknown length
07364 209043: contig of 1680 bp in length
09044 209143: gap of unknown length
09144 217567: contig of 8424 bp in length
17568 217667: gap of unknown length
17668 218485: contig of 818 bp in length.
```

1..218485
Location/Qualifiers
1..218485
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="RP43-145D13"
/clone_lib="RP43"
1..7548
/note="assembly_fragment
clone_end:SP6
vector_side:left"
7649..31917
/note="assembly_fragment"
32018..50433
/note="assembly_fragment"
50534..95274
/note="assembly_fragment"
95375..99988
/note="assembly_fragment"
100089..109873

misc_feature   109974..117619
                  /note="assembly_fragment"
misc_feature   117720..128625
                  /note="assembly_fragment"
misc_feature   128726..143521
                  /note="assembly_fragment"
misc_feature   143622..150832
                  /note="assembly_fragment"
misc_feature   150933..156073
                  /note="assembly_fragment"
misc_feature   156174..187450
                  /note="assembly_fragment"
misc_feature   187551..191471
                  /note="assembly_fragment"
misc_feature   191572..207263
                  /note="assembly_fragment"
misc_feature   207364..209043
                  /note="assembly_fragment"
misc_feature   209144..217567
                  /note="assembly_fragment"
misc_feature   217668..218485
                  /note="assembly_fragment
                  clone_end:T7
                  vector_side:right"

ORIGIN

Query Match   31.7%;   Score 435; DB 2;   Length 218485;
Best Local Similarity 99.3%;   Pred. No. 1.1e-230;
Matches 755;   Conservative 0;   Mismatches 4;   Indels 1; (

Qy   601 CAGGTGCACTTTGATGAGGGAAGGCTGCTCACCTGACGTGTTTCCATCCCACA
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 52585 CAGGTGCACTTTGATGAGGGAAGGCTGCTCACCTGACGTGTTTCCATCCCACA

Qy   661 GTGCTGGCCCTGCGCTGCCTGGAGAATTCTCAGCCACTGCGGCCAGTTCCCTCG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 52645 GTGCTGGCCCTGCGCTGCCTGGAGAATTCTCAGCCACTGCGGCCAGTTCCCTCG

Qy   721 CAGCTCCGCCTGCCAGGTGTCTGGCGTGTTGGCCCTGCCAGGGTCCTCCC
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 52705 CAGCTCCGCCTGCCAGGTGTCTGGCGTGTTGGCCCTGCCAGGGTCCTCCC

Qy   781 ATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACCTACTTCG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 52765 ATCCGCACCCTCCCCTGGGCCCATCTCAAGGCTGCCCCCTTCCTCACCTACTTCG

Qy   841 TTCCAGGTTCACTGAGGGGCCCTGGTCTCCCCACAGTCGTCCCAGGCTGCCGGCT
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 52825 TTCCAGGTTCACTGAGGGGCCCTGGTCTCCCCGCAGTCGTCCCAGGCTGCCGGCT

Qy   901 CGACAGCTCTCGGGCACCCGGTCCCCTGCCCCTCTGCCCCCTCAGCTGTCTCTTTGCT
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 52885 CGACAGCTCTCGGGCACCCGGTCCCCTGCCCCTCTGCCCCCTCAGCTGTCTCTTTGCT

Qy   961 CCTGCCCCTCCAGAGGCTGCCTGGGCCTGTTCACGTGTTTCCATCCCACA
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 52945 CCTGCCCCTCCAGAGGCTGCCTGGGCCTGTTCACGTGTTTCCATCCCACA

Qy 1021 ACAGTATTCCCACTCTTATCTTACAACT-CCCCACCGCCCACTCTCCACCTCAC
          ||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db 53005 ACAGTATTCCCACTCTTATCTTACAACTCCCCCACCGCCCACTCTCCACCTCAC

Qy 1080 CCCCAATCCCTGACCCTTTGAGGCCCCCAGTGATCTGGACTCCCCCCTGGCCACA
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 53065 CCCCAATCCCTGACCCTTTGAGGCCCCCAGTGATCTGGACTCCCCCCTGGCCACA

Qy 1140 CCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCAAGAGACCCCA
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 53125 CCAGGGCATTGTGTTCACTGTACTCTGTGGGCAAGGATGGGTCAAGAGACCCCA

Qy 1200 GGCACTAAGAGGGGCTGGACCTGGCGCAGGAAGCCAAAGAGACTGGGCCTAGGC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
         Sison,C., Stantripop,S., Thomas,J.W., Thomas,P.J., Tsipour
         Vogt,J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E
  TITLE    NISC Comparative Sequencing Initiative
  JOURNAL  Unpublished
REFERENCE   2  (bases 1 to 180222)
  AUTHORS  Green,E.D.
  TITLE    Direct Submission
  JOURNAL  Submitted (08-AUG-2002) NIH Intramural Sequencing Center,
           Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE   3  (bases 1 to 180222)
  AUTHORS  Green,E.D.
  TITLE    Direct Submission
  JOURNAL  Submitted (05-JUN-2003) NIH Intramural Sequencing Center,
           Grovemont Circle, Gaithersburg, MD 20877, USA
           On Jun 5, 2003 this sequence version replaced gi:25447561.
COMMENT    ------ Genome Center
           Center: NIH Intramural Sequencing Center
           Center code: NISC
           Web site: http://www.nisc.nih.gov
           Contact: nisc_zoo@nhgri.nih.gov
           ------ Project Information
           Center project name: dbg
           Center clone name: 436K21
           ------

           The sequence data in this record represents an 'enhanced'
           version of a Phase 2 submission. Specifically, the indicat
           order and orientation of each sequence contig has been
           established using one or more of the following: read-pair
           data from individual subclones, overlaps with neighboring
           clones, alignment with available reference sequence (e.g.,
           human), and/or confirmation by PCR testing. In addition,
           the sequence assembly is based on at least 8X average
           coverage in Q20 bases and has been reviewed to rule out
           gross misassemblies, the low-quality ends of sequence
           contigs have been trimmed away, and each base is associate
           with a Phrap-derived quality score.
           ------ Summary Statistics
           Sequencing vector: plasmid; n/a; 100% of reads
           Chemistry: Dye-terminator Big Dye; 100% of reads
           Assembly program: Phrap; version 0.990319
           Consensus quality: 178450 bases at least Q40
           Consensus quality: 179050 bases at least Q30
           Consensus quality: 179435 bases at least Q20
           Insert size: 169000; agarose-fp
           Insert size: 179622; sum-of-contigs
           Quality coverage: 12.21x in Q20 bases; agarose-fp
           Quality coverage: 11.49x in Q20 bases; sum-of-contig
           ------
         * NOTE: This is a 'working draft' sequence. It currently
         * consists of 7 contigs. Gaps between the contigs
         * are represented as runs of N. The order of the pieces
         * is believed to be correct as given, however the sizes
         * of the gaps between them are based on estimates that hav
         * provided by the submittor.
         * This sequence will be replaced
         * by the finished sequence as soon as it is available and
         * the accession number will be preserved.
         *      1     27615: contig of 27615 bp in length
         *  27616     27715: gap of unknown length
         *  27716     63581: contig of 35866 bp in length
         *  63582     63681: gap of unknown length
         *  63682     79016: contig of 15335 bp in length
         *  79017     79116: gap of unknown length
         *  79117     91018: contig of 11902 bp in length
         *  91019     91118: gap of unknown length
         *  91119     94289: contig of 3171 bp in length
         *  94290     94389: gap of unknown length
         *  94390    167164: contig of 72775 bp in length
         * 167165    167264: gap of unknown length
         * 167265    180222: contig of 12958 bp in length
FEATURES          Location/Qualifiers
     source        1..180222
                   /organism="Sus scrofa"
```

```
TTAAGAGGGGCTGGACCTGGCGGCAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGA 53244

CCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGA 1319
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCAAATGTGAGGGGCGAGAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGA 53304

TAAAACAGATATTATTTTTATTATTATTATTGTGACAAA 1359
|||||||||||||||||||||||||||||||||||||||
TAAAACAGATATTATTTTTATTATTATTATTGTGACAAA 53344

ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

r,M.J., King,G.E., Xu,J. and Secrist,H.
itions and methods for the therapy and diagnosis of colon

24        195 bp    DNA    linear    PAT 18-MAR-2002
ce 66 from Patent WO0196389.
24
24.1  GI:19574864

; WO 0196389-A 66 20-DEC-2001;
CORPORATION (US)
Location/Qualifiers
1..195
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

apiens (human)
apiens

arity    13.0%;  Score 179;  DB 6;  Length 195;
100.0%;  Pred. No. 6.6e-88;
onservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

GTGTTCACTGTACTCTGTGGGCAAGGATGGTCCAGAAGACCCCACTTCAGGCACT 1205
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTGTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAGACCCCACTTCAGGCACT 76

GGGGCTGACCTGGCGCGCAGGAAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCC 1265
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGGGCTGACCTGGCGCGCAGGAGCCAAAGAGACTGGGCCTAGGCCAGGAGTTCCC 136

GTGAGGGGCGAGAAAACAAGACAAGCTCCTCCCTGAGAATTCCCTGTGGATTTTT 1324
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTGAGGGGCGAGAAAACAAGACAAGCTCCTCCCTTGAGAATTCCCTGTGGATTTTT 195

92        180222 bp   DNA    linear    HTG 05-JUN-2003
rofa clone RP44-436K21, WORKING DRAFT SEQUENCE, 7 ordered
92
92.4  GI:31415895
TGS_PHASE2; HTGS_DRAFT.
rofa (pig)
rofa
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
ses 1 to 180222)

llis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
ley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Cariaga,K.,
, Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X.,
J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L, Hu,P.,
B., Idol,J.R., Karlins,E., Kwong,P., Laric,P., Lee-Lin,S.-Q.,
i,R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Masiello,C.,
i,B., McDowell,J., Paguirigan,C., Pearson,R., Portnoy,M.E.,
i,A., Reddix-Dugue,N., Schandler,K., Schueler,M.G., Shah,K.,
```

```
/mol_type="genomic DNA"
/db_xref="taxon:9823"
/clone="RP44-436K21"
/clone_lib="RP44"
1..27615
/note="assembly_fragment
clone_end:SP6
vector_side:left"
27716..63581
/note="assembly_fragment"
63682..79016
/note="assembly_fragment"
79117..91018
/note="assembly_fragment"
91119..94289
/note="assembly_fragment"
94390..167164
/note="assembly_fragment"
155535..180222
/note="clone overlaps with GenBank Accession Number
AC127472 clone RP44-398N15 (center project name dbh)"
167265..180222
/note="assembly_fragment
clone_end:T7
vector_side:right"
```

```
arity       10.6%;    Score 145;    DB 2;    Length 180222;
conservative      0;   Mismatches      0;   Indels      0;   Gaps      0;
```

```
GCCCGCCGGCTCCCCCCTCCCCCGATCCCTCGGGTCCCGGGATGGGGGGCGGGTGAGG 84
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCCCGCCGGCTCCCCCCTCCCCCGATCCCTCGGGTCCCGGGATGGGGGGCGGGTGAGG 95940
```

```
GCACAGCCCCCGGCCCCCATGGCCGCCCGTCGGAGCCAGAGGCGGAGGGGGCGCCGG 144
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCACAGCCCCCGCCCCCATGGCCGCCCGTCGGAGCCAGAGGCGGAGGGGGCGCCGG 96000
```

```
GAGCCGGGCACCGCCCTGCTGG 169
||||||||||||||||||||||
GAGCCGGGCACCGCCCTGCTGG 96025
```

```
195                 130254 bp    DNA      linear    HTG 06-JUN-2003
s norvegicus clone RP31-258K6 strain Brown Norway, WORKING
SEQUENCE, 12 ordered pieces.
195.
195.3    GI:31442440
HTGS_PHASE2; HTGS_DRAFT.
s norvegicus (Norway rat)
s norvegicus
yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
lia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
s.
ases 1 to 130254)
ellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
sley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Cariaga,K.,
., Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X.,
,J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
,B., Idol,J.R., Karlins,E., Kwong,P., Laric,P., Lee-Lin,S.-Q.,
pi,R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Masiello,C.,
ri,B., McDowell,J., Paguirigan,C., Pearson,R., Portnoy,M.E.,
d,A., Reddix-Dugue,N., Schandler,K., Schueler,M.G., Shah,K.,
,C., Stantripop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V.,
J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
Comparative Sequencing Initiative
lished
ases 1 to 130254)
,E.D.
t Submission
```

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicat
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8X average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associate
with a Phrap-derived quality score.
             --- Summary Statistics
             Sequencing vector: plasmid; n/a; 100% of reads
             Chemistry: Dye-terminator Big Dye; 100% of reads
             Assembly program: Phrap; version 0.990319
             Consensus quality: 128169 bases at least Q40
             Consensus quality: 128674 bases at least Q30
             Consensus quality: 128972 bases at least Q20
             Insert size: 150000; agarose-fp
             Insert size: 129154; sum-of-contigs
             Quality coverage: 10.78x in Q20 bases; agarose-fp
             Quality coverage: 12.52x in Q20 bases; sum-of-contig
             ----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that ha
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
```
*        1    10521: contig of 10521 bp in length
*    10522    10621: gap of unknown length
*    10622    13327: contig of 2706 bp in length
*    13328    13427: gap of unknown length
*    13428    28924: contig of 15497 bp in length
*    28925    29024: gap of unknown length
*    29025    39201: contig of 10177 bp in length
*    39202    39301: gap of unknown length
*    39302    41906: contig of 2605 bp in length
*    41907    42006: gap of unknown length
*    42007    70095: contig of 28089 bp in length
*    70096    70195: gap of unknown length
*    70196    77561: contig of 7366 bp in length
*    77562    77661: gap of unknown length
*    77662    94161: contig of 16500 bp in length
*    94162    94261: gap of unknown length
*    94262    97982: contig of 3721 bp in length
*    97983    98082: gap of unknown length
*    98083   105400: contig of 7318 bp in length
*   105401   105500: gap of unknown length
*   105501   127855: contig of 22355 bp in length
*   127856   127955: gap of unknown length
*   127956   130254: contig of 2299 bp in length.
```

Location/Qualifiers
1..130254
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="Brown Norway"
/db_xref="taxon:10116"
/clone="RP31-258K6"
/clone_lib="RP31"
1..10521
/note="assembly_fragment
clone_end:SP6
vector_side:left"
10622..13327
/note="assembly_fragment"
13428..28924
/note="assembly_fragment"
/note="assembly_fragment"
29025..39201
/note="assembly_fragment"
39302..41906
/note="assembly_fragment"
42007..70095
/note="assembly_fragment"
70196..77561
/note="assembly_fragment"
77662..94161
/note="assembly_fragment"
94262..97982
/note="assembly_fragment"
98083..105400
/note="assembly_fragment"
105501..127855
/note="assembly_fragment"
117804..130254
/note="clone overlaps with GenBank Accession Number
AC134317 clone RP31-421B16 (center project name dce)"
127796..130254
/note="assembly_fragment
missing T7 clone end on 3' end of insert"

rity   8.7%;   Score 119;   DB 2;   Length 130254;
nservative 100.0%;   Mismatches   0;   Indels   0;   Gaps   0;
　Pred. No. 3.9e-54;

```
TCGGGTCCCGGGATGGGGGGGGGGGGTGAGGCAGGCACAGCCCCCCGCCCCCATGGC 110
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCGGGTCCCGGGATGGGGGGGGGGGGTGAGGCAGGCACAGCCCCCCGCCCCCATGGC 64163

GTCGGAGCCAGAGGCGGAGGGGGCGGCGCCCGGGGGGGAGCCGGGCACCGGCCTGCTGG 169
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTCGGAGCCAGAGGCGGAGGGGGCGCGCCCGGGGGGGAGCCGGGCACCGGCCCTGCTGG 64222
```

5　165316 bp   DNA   linear   HTG 19-NOV-2002
norvegicus clone CH230-320N23, WORKING DRAFT SEQUENCE, 5
ed pieces.
5
5.4 GI:25100662
GS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
norvegicus (Norway rat)
norvegicus
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

es 1 to 165316)
Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
echi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,

Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Sou
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederic
Delgado,O., Denson,S., Derano,C., Ding,Y., Dinh,H., Divya,K
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Ea
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G
Fernandez,C.M., Finley,M., Flagg,N., Forbes,L., Foster,M., Fo
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Gar
Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Gue
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jac
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jol
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangu
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Na
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Park
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Re
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., S
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.
Sneed,A., Sodergren,E., Song,X.-z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Tayl
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,J., Usma
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., War
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wlecyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yo
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhou,S., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.
Weinstock,G. and Gibbs,R.A.

TITLE   Direct Submission
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 165316)
AUTHORS   Worley,K.C.
TITLE   Direct Submission
JOURNAL   Submitted (25-APR-2002) Human Genome Sequencing Center, Dep
          of Molecular and Human Genetics, Baylor College of Medicine
          Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 165316)
AUTHORS   Rat Genome Sequencing Consortium.
TITLE   Direct Submission
JOURNAL   Submitted (19-NOV-2002) Human Genome Sequencing Center, Dep
          of Molecular and Human Genetics, Baylor College of Medicine
          Baylor Plaza, Houston, TX 77030, USA
COMMENT   On Nov 19, 2002 this sequence version replaced gi:23616728.
          The sequence in this assembly is a combination of BAC based
          and whole genome shotgun sequencing reads assembled using
          (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig de
          in the feature table below represents a scaffold (a 'contig-scaffol
          assembly ('contig-scaffold'). Within each contig-scaffol
          individual sequence contigs are ordered and oriented, and s
          by sized gaps filled with Ns to the estimated size. The sec
          may extend beyond the ends of the clone and there may be se
          contigs within a contig-scaffold that consist entirely of
          genome shotgun sequence reads. Both end sequences and whole
          shotgun sequence only contigs will be indicated in the feat
          table.
          ------------- Genome Center
            Center: Baylor College of Medicine
            Center code: BCM

```
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murin
REFERENCE   1 (bases 1 to 203083)
  AUTHORS   Metzker,M.L., Lewis,L.R., Hume,J., Edwards,C., Harris,C.,
            Dederich,D., Thomas,S., Okwuonu,G., Carlock,C., Garner,T.,
            Addison,D., Pace,A., Williams,G., Bonnin,D., Brooks,A., Br
            Buhay,C., Bunac,C., Burkett,C., Chacko,J., Chen,G., Chen,Z
            Cox,C., Davis,C., Delgado,O., Ding,Y., Dugan-Rocha,S.,
            Fernandez,C., Ferraguto,D., Forcum-Tansey,J., Gill,R.,
            Gorrell,J.H., Gunaratne,P., Haller,G., Hernandez,J., Hogue
            Hosak,H., Hou,X., Huber,J., Jackson,L., Jia,Y., Kelly,J.,
            Kovar,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Marti
            Massey,E., McLeod,M.P., Mei,G., Moore,S., Morgan,M., Morri
            Neal,D., Nelson,A., Nguyen,R., Nguyen,N., Oguh,M., Parish,
            Perez,L., Reiter,D., Say,J., Shen,H., Vasquez,L., Watlingt
            Williamson,A., Wrensford,G., Zhou,X., Bouck,J., Hodgson,A.
            Muzny,D.M., Rives,M., Scherer,S., Sodergren,E., Weinstock,
            Worley,K. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 203083)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-MAY-2000) Human Genome Sequencing Center, De
            of Molecular and Human Genetics, Baylor College of Medicin
            Baylor Plaza, Houston, TX 77030, USA
            On Jun 25, 2001 this sequence version replaced gi:12621364
COMMENT     --------------- Genome Center
            Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu/
            Contact: hgsc-help@bcm.tmc.edu
            ---------------  Project Information
            Center project name: MAFO
            Center clone name: RP23-168P5
            ---------------  Summary Statistics
            Sequencing vector: M13; L08821
            Chemistry: Dye-primer Bodipy: 48% of reads
            Chemistry: Dye-terminator Big Dye: 52% of reads
            Assembly program: Phrap; version 0.990329
            Consensus quality: 212648 bases at least Q40
            Consensus quality: 218902 bases at least Q30
            Consensus quality: 222384 bases at least Q20
            Estimated insert size: 210656; sum-of-contigs estimati
            Quality coverage: 0x in Q20 bases; agarose-fp estimati
            Quality coverage: 7.2x in Q20 bases; sum-of-contigs es
          * NOTE: Estimated insert size may differ from sequence ler
          *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_da
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 7 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *        1    62152: contig of 62152 bp in length
          *    62153    62252: gap of unknown length
          *    62253   118772: contig of 56520 bp in length
          *   118773   118872: gap of unknown length
          *   118873   148924: contig of 30052 bp in length
          *   148925   149024: gap of unknown length
          *   149025   167231: contig of 18207 bp in length
          *   167232   167331: gap of unknown length
          *   167332   189907: contig of 22576 bp in length
          *   189908   190007: gap of unknown length
          *   190008   196537: contig of 6530 bp in length
          *   196538   196637: gap of unknown length
          *   196638   203083: contig of 6446 bp in length.
```

```
b site: http://www.hgsc.bcm.tmc.edu/
ntact: hgsc-help@bcm.tmc.edu
-------  Project Information
nter project name: GUXG
nter clone name: CH230-320N23
-------  Summary Statistics
sembly program: Phrap; version 0.990329
nsensus quality: 155971 bases at least Q40
nsensus quality: 157321 bases at least Q30
nsensus quality: 158253 bases at least Q20
timated insert size: 159662; sum-of-contigs estimation
ality coverage: 7x in Q20 bases; sum-of-contigs estimation
-------
: Estimated insert size may differ from sequence length
ee http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
: This is a 'working draft' sequence. It currently
ists of 5 contigs. The true order of the pieces
ot known and their order in this sequence record is
trary. Gaps between the contigs are represented as
of N, but the exact sizes of the gaps are unknown.
record will be updated with the finished sequence
on as it is available and the accession number will
reserved.
    1   159511: contig of 159511 bp in length
9512   159611: gap of unknown length
9612   160669: contig of 1058 bp in length
0670   160769: gap of unknown length
0770   161864: contig of 1095 bp in length
1865   161964: gap of unknown length
1965   163701: contig of 1737 bp in length
3702   163801: gap of unknown length
3802   165316: contig of 1515 bp in length.
Location/Qualifiers
1..165316
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-320N23"
1..2103
/note="wgs end extension
clone end:T7"
8259..9143
/note="clone_boundary
clone end:T7
site:
end_sequence:B2171094"
46341..159014
/note="clone_boundary
clone end:Sp6
site:
end_sequence:B2171096"
```

```
arity    8.7%;  Score 119;  Pred. No. 3.9e-54;  Length 165316;
onservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

CTCGGGTCCCGGGATGGGGGGCGGTGAGGGAGGCACAGCCCCCGCCCCCCATGGC 110
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTCGGGTCCCGGGATGGGGGGCGGGCGGTGAGGGCAGGCACAGCCCCCCATGGC 120224

CGTCGGAGCCAGAGGCCGGAGGGGGCGCCCGGGGGGGAGCCGGCACCGCCCTGCTGG 169
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
CGTCGGAGCCAGAGGCCGGAGGGGGGCGCGGGGGGGAGCCGGCACCGCCCTGCTGG 120165
```

```
159     203083 bp    DNA     linear   HTG 27-JUN-2001
usculus chromosome 11 clone RP23-168P5, WORKING DRAFT
CE, 7 unordered pieces.
59.23  GI:14547768
```

F

Location/Qualifiers
1..203083
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-168P5"

rity   8.7%;   Score 119;   DB 2;   Length 203083;
100.0%;   Pred. No. 4e-54;
nservative 0; Mismatches 0; Indels 0; Gaps 0;

TCGGGTCCCCGGGATGGGGGGCGGTGAGGCAGGCACAGCCCCCGCCCCCATGGC 110
||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCGGGTCCCCGGGATGGGGGGCGGTGAGGCAGGCACAGCCCCCGCCCCCATGGC 51674

GTCGGAGCCAGAGGCGGAGGGGGGCCGGGGGAGCCGGCACCGCCCTGCTGG 169
||||||||||||||||||||||||||||||||||||||||||||||||||||
GTCGGAGCCAGAGGCGGAGGGGGGCCGGGGGAGCCGGCACCGCCCTGCTGG 51615

23   223877 bp   DNA   linear   HTG 10-MAY-2003
norvegicus clone CH230-154B15, WORKING DRAFT SEQUENCE, 3
ed pieces.

3.8 GI:30521223
GS PHASE1; HTGS DRAFT; HTGS_FULLTOP.   •
norvegicus (Norway rat)
norvegicus
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

es 1 to 223877)
O.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
, Allen,H., Alsbrooks,S., Amin,A., Arguiano,D.,
echi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
as,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chu,J.,
nd,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
.Escotto.M. Eugene,C., Evans,C.A., Falls,T., Fan,G.,
lez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
eorgis,E. Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
ne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
lez,R. Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
n,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
ry,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
, Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
, Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
shewa,L., Louiseged,H., Lozado,R.J., Lu,X., Ma,J.,
wari,M., Mahindartie,M., Mahmoud,M., Malloy,K., Mangum,A.,
,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
ey,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
vljevic,A., Miner,G., Minja,B., Montemayor,J., Moore,S.,
,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
vis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
lemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
oak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
r,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.L.,
, Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,

Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Tay
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usm
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wa
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wlieczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yo
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H
Weinstock,G. and Gibbs,R.A.

TITLE   Direct Submission
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 223877)
AUTHORS   Worley,K.C.
TITLE   Direct Submission
JOURNAL   Submitted (06-NOV-2001) Human Genome Sequencing Center, De
           of Molecular and Human Genetics, Baylor College of Medicin
           Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 223877)
AUTHORS   Rat Genome Sequencing Consortium.
TITLE   Direct Submission
JOURNAL   Submitted (10-MAY-2003) Human Genome Sequencing Center, De
           of Molecular and Human Genetics, Baylor College of Medicin
           Baylor Plaza, Houston, TX 77030, USA
COMMENT   On May 10, 2003 this sequence version replaced gi:25008075
           The sequence in this assembly is a combination of BAC base
           and whole genome shotgun sequencing reads assembled using
           (http://www.hgsc.bcm.tmc.edu/projects/rat/).Each contig d
           in the feature table below represents a scaffold in the At
           assembly (a 'contig-scaffold'). Within each contig-scaffol
           individual sequence contigs are ordered and oriented, and
           by sized gaps filled with Ns to the estimated size. The se
           may extend beyond the ends of the clone and there may be s
           contigs within a contig-scaffold that consist entirely of
           genome shotgun sequence reads. Both end sequences and whol
           shotgun sequence only contigs will be indicated in the fea
           table.
           ----- Genome Center
                 Center: Baylor College of Medicine
                 Center code: BCM
                 Web site: http://www.hgsc.bcm.tmc.edu/
                 Contact: hgsc-help@bcm.tmc.edu
           ----- Project Information
                 Center project name: G10K
                 Center clone name: CH230-154B15
           ----- Summary Statistics
                 Assembly program: Atlas 3.0;
                 Consensus quality: 214785 bases at least Q40
                 Consensus quality: 216908 bases at least Q30
                 Consensus quality: 218593 bases at least Q20
                 Estimated insert size: 227169; sum-of-contigs estimati
                 Quality coverage: 7x in Q20 bases; sum-of-contigs esti
           * NOTE: Estimated insert size may differ from sequence len
           *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_da
           * NOTE: This is a 'working draft' sequence. It currently_
           *   consists of 3 contigs. The true order of the pieces
           *   is not known and their order in this sequence record is
           *   arbitrary. Gaps between the contigs are represented as
           *   runs of N, but the exact sizes of the gaps are unknown.
           *   This record will be updated with the finished sequence
           *   as soon as it is available and the accession number will
           *   be preserved.
           *   1   221327: contig of 221327 bp in length
           *   221328 221427: gap of unknown length
           *   221428 222652: contig of 1225 bp in length
           *   222653 222752: gap of unknown length
           *   222753 223877: contig of 1125 bp in length.
FEATURES   Location/Qualifiers
source   1..223877
           /organism="Rattus norvegicus"

F

```
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Rigc
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,I
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,A., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Tay
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usn
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wa
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yc
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,F
Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 225077)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-NOV-2002) Human Genome Sequencing Center, De
            of Molecular and Human Genetics, Baylor College of Medici
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3 (bases 1 to 225077)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAY-2003) Human Genome Sequencing Center, De
            of Molecular and Human Genetics, Baylor College of Medici
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 10, 2003 this sequence version replaced gi:2518870(
            The sequence in this assembly is a combination of BAC base
            and whole genome shotgun sequencing reads assembled using
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig (
            in the feature table below represents a scaffold in the A
            assembly (a 'contig-scaffold'). Within each contig-scaffo
            individual sequence contigs are ordered and oriented, and
            by sized gaps filled with Ns to the estimated size. The s
            may extend beyond the ends of the clone and there may be
            contigs within a contig-scaffold that consist entirely of
            genome shotgun sequence reads. Both end sequences and who
            shotgun sequence only contigs will be indicated in the fe
            table.
            ----- Genome Center
              Center: Baylor College of Medicine
              Center code: BCM
              Web site: http://www.hgsc.bcm.tmc.edu/
              Contact: hgsc-help@bcm.tmc.edu
            ----- Project Information
              Center project name: GHDQ
              Center clone name: CH230-46E21
            ----- Summary Statistics
              Assembly program: Atlas 3.0;
              Consensus quality: 214599 bases at least Q40
              Consensus quality: 216978 bases at least Q30
              Consensus quality: 218480 bases at least Q20
              Estimated insert size: 225046; sum-of-contigs estimat
              Quality coverage: 8x in Q20 bases; sum-of-contigs est
          *  NOTE: Estimated insert size may differ from sequence le
          *    (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_d
          *  NOTE: This is a 'working draft' sequence. It currently
          *    consists of 1 contigs. Gaps between the contigs
          *    are represented as runs of N. The order of the pieces
          *    is believed to be correct as given, however the sizes
          *    of the gaps between them are based on estimates that ha
          *    provided by the submittor.
          *  This sequence will be replaced
          *  by the finished sequence as soon as it is available and
          *  the accession number will be preserved.
          *
FEATURES          Location/Qualifiers
     source       1..225077
                  /organism="Rattus norvegicus"
                  /mol_type="genomic DNA"
```

```
                  /mol_type="genomic DNA"
                  /db_xref="taxon:10116"
                  /clone="CH230-154B15"
                  1..1822
                  /note="wgs contig"
                  complement(217607..218056)
                  /note="clone_boundary
                  clone end:T7_
                  site:EcoRI
                  end_sequence:RWBBO08TJB"

arity      8.7%;  Score 119;  DB 2;  Length 223877;
   100.0%;  Pred. No. 4e-54;  Indels  0;  Gaps  0;
onservative  0; Mismatches  0;

CTCGGGTCCCGGGATGGGGGGGGTGAGGCAGGCACAGCCCCCGCCCCATGGC  110
||||||||||||||||||||||| ||||||||||||||||||||||||||||
CTCGGGTCCCGGGATGGGGGGGGCGGTGAGGCAGGCACAGCCCCCGCCCCATGGC  139398

CGTCGGAGCCAGAGGCGGAGGGGGCGCCCGGGGGAGCCGGCACCGGCCCTGCTGG  169
||||||||||||||||||||||||||||||||||||||||||||||||||||||
CGTCGGAGCCAGAGGCGGAGGGGGCGCCCGGGGGAGCCGGCACCGGCCCTGCTGG  139457
```

```
...63        225077 bp    DNA        linear    HTG 10-MAY-2003
...63  norvegicus clone CH230-46E21, WORKING DRAFT SEQUENCE.
...63
...63.3    GI:30521905
HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
  norvegicus (Norway rat)
  norvegicus
...ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
...ia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
...s.
ases 1 to 225077)
.D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
.C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
abechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
in,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
lo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
t,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
nas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
o,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
land,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
a,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
do,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
r,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
ndez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
r,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
georgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
atne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
y,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
ndez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
ns,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
on,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
thy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
.C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
.., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
suhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
hwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
m,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
ney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
avljevic,A., Miner,G., Minja,S., Montemayor,J., Moore,S.,
n,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
rvis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
elemeh,O., Okwuonu,G., Olaznpunsagoon,A., Pal,S., Parks,K.,
rnak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
er,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
```

```
Query Match          8.7%;  Score 119;  DB 10;   Length 234182;
Best Local Similarity 100.0%;  Pred. No. 4.1e-54;
Matches 119;  Conservative 0;  Mismatches 0;  Indels 0;  G

Qy    51 ATCCCTCGGGTCCCGGGATGGGGGGCGGGGGGCGGTGAGGCAGGCACAGCCCCCGCCCCCA
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 75038 ATCCCTCGGGTCCCGGGATGGGGGGCGGGGGGCGGTGAGGCAGGCACAGCCCCCGCCCCCA

Qy   111 CGCCCGTCGGAGCCAGCCGGAGGCGGAGGGGCCCGGGGGGGAGCCGGGGACCGCCCCTGC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 74978 CGCCCGTCGGAGCCAGCCGGAGGCGGAGGGGCCCGGGGGGGAGCCGGGGACCGCCCCTGC
```

```
RESULT 22
AC126237
LOCUS       AC126237         212093 bp    DNA    linear    HTG 06-
DEFINITION  Canis familiaris clone RP81-414O22, WORKING DRAFT SEQUENCE
            ordered pieces.
ACCESSION   AC126237
VERSION     AC126237.5  GI:31442445
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      Canis familiaris (dog)
ORGANISM    Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1 (bases 1 to 212093)
AUTHORS     Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjami
            Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Car
            Chu,G., Coleman,B., Coleman,H., Engle,J., Granite,S., Guan
            Gupta,J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
            Hurle,B., Idol,J.R., Karlins,E., Kwong,P., Laric,P., Lee-L
            Legaspi,R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Masi
            Maskeri,B., McDowell,J., Paguirigan,C., Pearson,R., Portno
            Prasad,A., Reddix-Dugue,N., Schandler,K., Schueler,M.G., S
            Sison,C., Stantripop,S., Thomas,J.W., Thomas,P.J., Tsipour
            Vogt,J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E
            NISC Comparative Sequencing Initiative
TITLE       Unpublished
JOURNAL     2 (bases 1 to 212093)
REFERENCE   Green,E.D.
AUTHORS     Direct Submission
TITLE       Submitted (04-JUL-2002) NIH Intramural Sequencing Center,
JOURNAL     Grovemont Circle, Gaithersburg, MD 20877, USA
            3 (bases 1 to 212093)
REFERENCE   Green,E.D.
AUTHORS     Direct Submission
TITLE       Submitted (06-JUN-2003) NIH Intramural Sequencing Center,
JOURNAL     Grovemont Circle, Gaithersburg, MD 20877, USA
            On Jun 6, 2003 this sequence version replaced gi:27476131.
COMMENT     ---------------- Genome Center
                  Center: NIH Intramural Sequencing Center
                  Center code: NISC
                  Web site: http://www.nisc.nih.gov
                  Contact: nisc_zoo@nhgri.nih.gov
            ---------------- Project Information
                  Center project name: cwg
                  Center clone name: 414O22
            ---------------- Summary Statistics
                  Sequencing vector: plasmid; n/a; 100% of reads
```

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicat
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8X average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associate
with a Phrap-derived quality score.

```
/db_xref="taxon:10116"
/clone="CH230-46B21"
1..1402
/note="wgs_contig"
```

```
urity   100.0%;  Pred. No. 4e-54;
nservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
          8.7%;  Score 119;  DB 2;   Length 225077;

TCGGGTCCCGGGATGGGGGGCGGTGAGGCAGGCAGCCCCCGCCCCCATGGC 110
|||||||||||||||||||||||||||||||||||||||||||||||||||
TCGGGTCCCGGGATGGGGGGCGGTGAGGCAGGCAGCCCCCGCCCCCATGGC 199986

GTCGGAGCCAGCCGGAGGCGGCGCCCGGGGGGGAGCCGGGGCACCGCCCTGCTGG 169
||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTCGGAGCCAGCCGGAGGCGGCGCCCGGGGGGGAGCCGGGGCACCGCCCTGCTGG 199927
```

```
07           234182 bp    DNA    linear    ROD 17-NOV-2001
DNA sequence from clone RP23-422L16 on chromosome 11,
e sequence.
07
07.5  GI:17017790
culus (house mouse)
culus
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ses 1 to 234182)
A.
Submission
ed (17-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
geshire, CB10 1SA, UK. E-mail enquiries:
ry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
20, 2001 this sequence version replaced gi:16605765.
sequence assembly data is compared from overlapping clones.
ifferences are found these are annotated as variations
er with a note of the overlapping clone name. Note that the
ion annotation may not be found in the sequence submission
ponding to the overlapping clone, as we submit sequences with
small overlap as described above.
equence was finished as follows unless otherwise noted: all
s were either double-stranded or sequenced with an alternate
ry or covered by high quality data (i.e., phred quality >=
n attempt was made to resolve all sequencing problems, such
pressions and repeats; all regions were covered by at least
asmid subclone or more than one M13 subclone; and the
ly was confirmed by restriction digest. The following
iations are used to associate primary accession numbers given
feature table with their source databases: Em:, EMBL; Sw:,
ROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
se can be found at
/www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-422L16 is
ne RPCI-23 Mouse PAC Library
cted by the group of Pieter de Jong.
: pBACe3.6
equence is the entire insert of clone RP23-422L16.
ther details see http://www.chori.org/bacpac/home.htm
Location/Qualifiers
1..234182
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-422L16"
/clone_lib="RPCI-23"
complement(84050..84131)
/note="Sequence from uni-directional primer reads and dGTP
big dye terminator reads only."
```

Chemistry: Dye-terminator Big Dye; 100% of reads
assembly program: Phrap; version 0.990319
Consensus quality: 208984 bases at least Q40
Consensus quality: 209673 bases at least Q30
Consensus quality: 210263 bases at least Q20
Insert size: 196000; agarose-fp
Insert size: 210693; sum-of-contigs
Quality coverage: 16.73x in Q20 bases; agarose-fp
Quality coverage: 15.57x in Q20 bases; sum-of-contigs

E: This is a 'working draft' sequence. It currently
sists of 15 contigs. Gaps between the contigs
represented as runs of N. The order of the pieces
believed to be correct as given, however the sizes
the gaps between them are based on estimates that have
vided by the submittor.
a sequence will be replaced
the finished sequence as soon as it is available and
accession number will be preserved.

```
1       3484:  contig of 3484 bp in length
3485    3584:  gap of unknown length
3585   28309:  contig of 24725 bp in length
28310  28409:  gap of unknown length
28410  38900:  contig of 10491 bp in length
38901  39000:  gap of unknown length
39001  43694:  contig of 4694 bp in length
43695  43794:  gap of unknown length
43795  57955:  contig of 14161 bp in length
57956  58055:  gap of unknown length
58056 137791:  contig of 79736 bp in length
137791 137891: gap of unknown length
137892 146862: contig of 8971 bp in length
146862 146962: gap of unknown length
146963 152584: contig of 5622 bp in length
152584 152684: gap of unknown length
152685 155668: contig of 2984 bp in length
155668 155768: gap of unknown length
155769 176557: contig of 20789 bp in length
176557 176657: gap of unknown length
176658 184091: contig of 7434 bp in length
184091 184191: gap of unknown length
184192 195037: contig of 10846 bp in length
195037 195137: gap of unknown length
195138 200818: contig of 5681 bp in length
200818 200918: gap of unknown length
200919 204476: contig of 3558 bp in length
204476 204576: gap of unknown length
204577 212093: contig of 7517 bp in length.
```

Location/Qualifiers
1..212093
/organism="Canis familiaris"
/mol_type="genomic DNA"
/db_xref="taxon:9615"
/clone="RP81-414O22"
/clone lib="RP81"
1..3484
/note="assembly_fragment
clone end:SP6
vector_side:left"
3585..28309
/note="assembly_fragment"
28410..38900
/note="assembly_fragment"
39001..43694
/note="assembly_fragment"
43795..57955
/note="assembly_fragment"
58056..137791
/note="assembly_fragment"
126505..212093
/note="clone overlaps with GenBank Accession Number
AC126925 clone RP81-332E11 (center project name cwp)"
137892..146862

misc_feature   /note="assembly_fragment"
                146963..152584
misc_feature   /note="assembly_fragment"
                152685..155668
misc_feature   /note="assembly_fragment"
                155769..176557
misc_feature   /note="assembly_fragment"
                176658..184091
misc_feature   /note="assembly_fragment"
                184192..195037
misc_feature   /note="assembly_fragment"
                195138..200818
misc_feature   /note="assembly_fragment"
                200919..204476
misc_feature   /note="assembly_fragment"
                204577..212093
misc_feature   /note="assembly_fragment
                clone_end:T7
                vector_side:right"

ORIGIN

Query Match   4.7%;   Score 65;   DB 2;   Length 212093;
Best Local Similarity   100.0%;   Pred. No. 6e-24;
Matches   65;   Conservative   0;   Mismatches   0;   Indels   0;

```
Qy     85 CAGGCACAGCCCCCGCCCCCATGGCCCGTCGGAGCAGGAGGCGGAGGGGGC
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 176827 CAGGCACAGCCCCCGCCCCCATGGCCCGTCGGAGCAGGAGGCGGAGGGGGC

Qy    145 GGGGA 149
          |||||
Db 176887 GGGGA 176891
```

RESULT 23
AF030100
LOCUS   AF030100   1239 bp   mRNA   linear   ROD 3
DEFINITION Mus musculus TWEAK mRNA, complete cds.
ACCESSION   AF030100
VERSION   AF030100.2 GI:33348855
KEYWORDS   .
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutel
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Muri
REFERENCE   1 (bases 1 to 1239)
  AUTHORS   Chicheportiche,Y., Bourdon,P.R., Xu,H., Hsu,Y.M., Scott,H
            Hession,C., Garcia,I. and Browning,J.L.
  TITLE   TWEAK, a new secreted ligand in the tumor necrosis factor
            that weakly induces apoptosis
  JOURNAL   J. Biol. Chem. 272 (51), 32401-32410 (1997)
  MEDLINE   98070415
   PUBMED   9405449
REFERENCE   2 (bases 1 to 1239)
  AUTHORS   Chicheportiche,Y., Bixler,S., Tizard,R. and Browning,J.
  TITLE   Direct Submission
  JOURNAL   Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridg
            Cambridge, MA 02142, USA
REFERENCE   3 (bases 1 to 1239)
  AUTHORS   Chicheportiche,Y., Bixler,S., Tizard,R. and Browning,J.
  TITLE   Direct Submission
  JOURNAL   Submitted (30-JUL-2003) Cell Biology, Biogen, 12 Cambridg
            Cambridge, MA 02142, USA
  REMARK   Sequence update by submitter
COMMENT   On Jul 30, 2003 this sequence version replaced gi:2707220
FEATURES   Location/Qualifiers
     source   1..1239
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /cell type="peritoneal macrophages"
     CDS   1..750
                     /note="ligand in the TNF family; secreted protei

/codon_start=1
/product="TWEAK"
/protein_id="AAC53517.2"
/db_xref="GI:33348856"
/translation="MAARRSQRRRGRRGEPGTALLAPIVLSLGLALACLGLLLVVVSL
GSWATLSAQEPSQEELTAEDRREPPELNPQTEESQDVVPFLEQLVRPRRSAPKGRKAR
PRRAIAAHYEVHRPRGQDGAQAGVDGTVSGWEETKINSSSPLRYDRQIGEFTVIRAGL
YYLYCQVHFDEGKAVYLKIDLLVNGVLALRCLEEFSATAASSPGPQLRLCQVSGLLPL
RPGSSLRIRTLPWAHLKAAPFLIYFGLFQVH"

arity   4.7%;   Score 64;   DB 10;   Length 1239;
onservative   0;   Mismatches   0;   Indels   0;   Gaps   0;
Pred. No. 1.4e-23;
arity 100.0%;

CCGCCCGTCGGAGCCAGAGGCGGAGGGCGGGCCCGGGGGGAGCCGGGCACCGCCCTG 165
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCGCCCGTCGGAGCCAGAGGCGGAGGGCGGGCCCGGGGGGAGCCGGGCACCGCCCTG 60

169

64

21   148555 bp   DNA   linear   HTG 06-JUN-2003
urus clone RP42-45D24, WORKING DRAFT SEQUENCE, 11 ordered

21
21.3   GI:31442443
TGS_PHASE2; HTGS_DRAFT.
urus (cow)
urus
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
e; Bovinae; Bos.
ses 1 to 148555)
llis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
ley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Cariaga,K.,
, Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X.,
J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
B., Idol,J.R., Karlins,E., Kwong,P., Laric,P., Lee-Lin,S.-Q.,
i,R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Masiello,C.,
i,B., McDowell,J., Paguirigan,C., Pearson,R., Portnoy,M.E.,
A., Reddix-Dugue,N., Schandler,K., Schueler,M.G., Shah,K.,
C., Stantripop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V.,
.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
omparative Sequencing Initiative
ished
ses 1 to 148555)
E.D.
Submission
ted (10-JUL-2002) NIH Intramural Sequencing Center, 8717
ont Circle, Gaithersburg, MD 20877, USA
ses 1 to 148555)
E.D.
Submission
ted (06-JUN-2003) NIH Intramural Sequencing Center, 8717
ont Circle, Gaithersburg, MD 20877, USA
6, 2003 this sequence version replaced gi:24137427.
------ Genome Center
enter: NIH Intramural Sequencing Center
enter code: NISC
eb site: http://www.nisc.nih.gov
ontact: nisc_zoo@nhgri.nih.gov
------ Project Information
enter project name: ddi
enter clone name: 045D24

quence data in this record represents an 'enhanced'
n of a Phase 2 submission. Specifically, the indicated
and orientation of each sequence contig has been

established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8X average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associate
with a Phrap-derived quality score.
------ Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 144765 bases at least Q40
Consensus quality: 146165 bases at least Q30
Consensus quality: 146960 bases at least Q20
Insert size: 151000; agarose-fp
Insert size: 147555; sum-of-contigs
Quality coverage: 8.83x in Q20 bases; agarose-fp
Quality coverage: 9.03x in Q20 bases; sum-of-contigs

------
* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that hav
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*   1   41911: contig of 41911 bp in length
*   41912   42011: gap of unknown length
*   42012   58977: contig of 16967 bp in length
*   58978   59078: gap of unknown length
*   59079   65825: contig of 6747 bp in length
*   65826   65925: gap of unknown length
*   65926   73946: contig of 8021 bp in length
*   73947   74046: gap of unknown length
*   74047   89416: contig of 15370 bp in length
*   89417   89516: gap of unknown length
*   89517   109889: contig of 20373 bp in length
*   109890   109989: gap of unknown length
*   109990   126878: contig of 16889 bp in length
*   126879   126978: gap of unknown length
*   126979   129099: contig of 2121 bp in length
*   129100   129199: gap of unknown length
*   129200   135799: contig of 6600 bp in length
*   135800   135899: gap of unknown length
*   135900   145102: contig of 9203 bp in length
*   145103   145202: gap of unknown length
*   145203   148555: contig of 3353 bp in length.

FEATURES   Location/Qualifiers
source   1..148555
            /organism="Bos taurus"
            /mol_type="genomic DNA"
            /db_xref="taxon:9913"
            /clone="RP42-45D24"
            /clone_lib="RP42"
misc_feature   1..41911
            /note="assembly_fragment
            clone_end:SP6
            vector_side:left"
misc_feature   42012..58978
            /note="assembly_fragment"
misc_feature   59079..65825
            /note="assembly_fragment"
misc_feature   65926..73946
            /note="assembly_fragment"
misc_feature   74047..89416
            /note="assembly_fragment"
misc_feature   89517..109889
            /note="assembly_fragment"
misc_feature   109990..126878

established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8X average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associat
with a Phrap-derived quality score.
------ Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 173760 bases at least Q40
Consensus quality: 174423 bases at least Q30
Consensus quality: 174916 bases at least Q20
Insert size: 152000; agarose-fp
Insert size: 175158; sum-of-contigs
Quality coverage: 17.46x in Q20 bases; agarose-fp
Quality coverage: 15.15x in Q20 bases; sum-of-contig

----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that ha
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
*   1   11425: contig of 11425 bp in length
*   11426   11525: gap of unknown length
*   11526   27554: contig of 16029 bp in length
*   27555   27654: gap of unknown length
*   27655   29974: contig of 2320 bp in length
*   29975   30074: gap of unknown length
*   30075   36241: contig of 6167 bp in length
*   36242   36341: gap of unknown length
*   36342   77002: contig of 40661 bp in length
*   77003   77102: gap of unknown length
*   77103 117018: contig of 39916 bp in length
*   117019 117118: gap of unknown length
*   117119 119041: contig of 1923 bp in length
*   119042 119141: gap of unknown length
*   119142 158388: contig of 39247 bp in length
*   158389 158488: gap of unknown length
*   158489 168033: contig of 9545 bp in length
*   168034 168133: gap of unknown length
*   168134 170716: contig of 2583 bp in length
*   170717 170816: gap of unknown length
*   170817 174429: contig of 3613 bp in length
*   174430 174529: gap of unknown length
*   174530 176258: contig of 1729 bp in length.

FEATURES   Location/Qualifiers
     source   1..176258
                     /organism="Canis familiaris"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9615"
                     /clone="RP81-332E11"
                     /clone_lib="RP81"
     misc_feature   1..88033
                     /note="clone overlaps with GenBank Accession Num
                     AC126237 clone RP81-414O22 (center project name
     misc_feature   1..11425
                     /note="assembly_fragment"
                     clone_end:SP6
                     vector_side:left"
     misc_feature   11526..27554
                     /note="assembly_fragment"
     misc_feature   27655..29974
                     /note="assembly_fragment"
     misc_feature   30075..36241
                     /note="assembly_fragment"

/note="assembly_fragment"
126979..148555
/note="clone overlaps with GenBank Accession Number
AC134961 clone RP42-406J16 (center project name djv)"
126979..129099
/note="assembly_fragment"
129200..135799
/note="assembly_fragment"
135900..145102
/note="assembly_fragment"
145203..148555
/note="assembly_fragment
clone_end:T7
vector_side:right"

larity   4.2%;   Score 58; DB 2; Length 148555;
Conservative 0; Mismatches 0; Pred. No. 4.7e-20; Indels 0; Gaps 0;

GTGCACTTGATGAGGGAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGGATG 658
||||||||||||||||||||||||||||||||||||||||||||||||||||
GTGCACTTGATGAGGGAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGGATG 60929

925   176258 bp   DNA   linear   HTG 06-JUN-2003
familiaris clone RP81-332E11, WORKING DRAFT SEQUENCE, 12
ed pieces.
925
925.6 GI:31442444
HTGS_PHASE2; HTGS_DRAFT.
familiaris (dog)
familiaris
yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
lia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
ases 1 to 176258)
ellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
sley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Cariaga,K.,
., Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X.,
J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
,B., Idol,J.R., Karlins,E., Kwong,N., Laric,P., Lee-Lin,S.-Q.,
pi,R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Masiello,C.,
ri,B., McDowell,J., Paguirigan,C., Pearson,R., Portnoy,M.E.,
d,A., Reddix-Duge,N., Schandler,K., Schueler,M.G., Shah,K.,
C., Stantripop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V.,
J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
Comparative Sequencing Initiative
lished
ases 1 to 176258)
,E.D.
Submission
tted (10-JUL-2002) NIH Intramural Sequencing Center, 8717
mont Circle, Gaithersburg, MD 20877, USA
ases 1 to 176258)
,E.D.
t Submission
tted (06-JUN-2003) NIH Intramural Sequencing Center, 8717
mont Circle, Gaithersburg, MD 20877, USA
n 6, 2003 this sequence version replaced gi:28209436.
------- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoo@nhgri.nih.gov
------- Project Information
Center project name: cwp
Center clone name: 332E11

equence data in this record represents an 'enhanced'
on of a Phase 2 submission. Specifically, the indicated
and orientation of each sequence contig has been

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicat
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8X average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associate
with a Phrap-derived quality score.
------------- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 148041 bases at least Q40
Consensus quality: 148385 bases at least Q30
Consensus quality: 148558 bases at least Q20
Insert size: 126000; agarose-fp
Insert size: 148636; sum-of-contigs
Quality coverage: 15.56x in Q20 bases; agarose-fp
Quality coverage: 13.19x in Q20 bases; sum-of-contigs
----

* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that hav
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

* 1   483: contig of 483 bp in length
* 484   583: gap of unknown length
* 584   19266: contig of 18683 bp in length
* 19267 19366: gap of unknown length
* 19367 21971: contig of 2605 bp in length
* 21972 22071: gap of unknown length
* 22072 24316: contig of 2245 bp in length
* 24317 24416: gap of unknown length
* 24417 36979: contig of 12563 bp in length
* 36980 37079: gap of unknown length
* 37080 42619: contig of 5540 bp in length
* 42620 42719: gap of unknown length
* 42720 46464: contig of 3745 bp in length
* 46465 46564: gap of unknown length
* 46565 61551: contig of 14987 bp in length
* 61552 61651: gap of unknown length
* 61652 99225: contig of 37574 bp in length
* 99226 99325: gap of unknown length
* 99326 110459: contig of 11134 bp in length
* 110460 110559: gap of unknown length
* 110560 147587: contig of 37028 bp in length
* 147588 147687: gap of unknown length
* 147688 149736: contig of 2049 bp in length.

FEATURES   Location/Qualifiers
     source   1..149736
                  /organism="Felis catus"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:9685"
                  /clone="RP86-474H17"
                  /clone_lib="RP86"
     misc_feature 1..483
                  /note="assembly_fragment
                  clone_end:T7
                  vector_side:left"
     misc_feature 584..19266
                  /note="assembly_fragment"
     misc_feature 19367..21971
                  /note="assembly_fragment"
     misc_feature 22072..24316

36342..77002
/note="assembly_fragment"
77103..117018
/note="assembly_fragment"
117119..119041
/note="assembly_fragment"
119142..158388
/note="assembly_fragment"
158489..168033
/note="assembly_fragment"
168134..170716
/note="assembly_fragment"
170817..174429
/note="assembly_fragment"
174530..176258
/note="assembly_fragment
clone_end:T7
vector_side:right"

arity   4.2%;   Score 58;   DB 2;   Length 176258;
Pred. No. 4.8e-20;   Indels   0;   Gaps   0;
nservative   0;   Mismatches   0;

TGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGGATG 658
|||||||||||||||||||||||||||||||||||||||||||||||||||||
TGCACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGGACTTGCTGGTGGATG 59938

39   149736 bp   DNA   linear   HTG 06-JUN-2003
catus clone RP86-474H17, WORKING DRAFT SEQUENCE, 12 ordered
.
39.5   GI:31442446
TGS_PHASE2; HTGS_DRAFT.
catus (cat)
catus
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
ses 1 to 149736)
llis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
ley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Cariaga,K.,
, Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X.,
J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
B., Idol,J.R., Karlins,E., Kwong,P., Laric,P., Lee-Lin,S.-Q.,
i.R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Masiello,C.,
i.B., McDowell,J., Paguirigan,C., Pearson,R., Portnoy,M.E.,
A., Reddix-Dugue,N., Schandler,K., Schueler,M.G., Shah,K.,
C., Stantripop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V.,
.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
omparative Sequencing Initiative
ished
ses 1 to 149736)
E.D.
Submission
ted (04-JUL-2002) NIH Intramural Sequencing Center, 8717
ont Circle, Gaithersburg, MD 20877, USA
ses 1 to 149736)
E.D.
Submission
ted (06-JUN-2003) NIH Intramural Sequencing Center, 8717
ont Circle, Gaithersburg, MD 20877, USA
.6, 2003 this sequence version replaced gi:26801264.
-------- Genome Center
enter: NIH Intramural Sequencing Center
enter code: NISC
eb site: http://www.nisc.nih.gov
ontact: nisc_zoo@nhgri.nih.gov
------- Project Information
enter project name: daa
enter clone name: 474H17

```
REFERENCE    unclassified.
             1 (bases 1 to 50)
AUTHORS      Ashkenazi,A.J., Marsters,S.A. and Pitti,R.
TITLE        Apo-3 ligand polypeptide
JOURNAL      Patent: JP 2001522584-A 3 20-NOV-2001;
             GENENTECH INC
COMMENT      OS   Unknown
             PN   JP 2001522584-A/3
             PD   20-NOV-2001
             PF   09-OCT-1998 JP 200051042
             PR   10-OCT-1997 US   60/062037,17-DEC-1997 US    60/069
             AVI  J ASHKENAZI,SCOT A MARSTERS,ROBERT PITTI
             PC   C12N15/09,A61K38/00,C07K14/705,C07K16/24,C12N15/00,A6
             CC   Description of Unknown Organism:Unknown
             FH   Key              Location/Qualifiers
             FT   source           1..50
             FT                    /organism='Unknown'.

FEATURES             Location/Qualifiers
     source          1..50
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"

ORIGIN

Query Match           3.6%;  Score 50;  DB 6;  Length 50;
Best Local Similarity  100.0%;  Pred. No. 7e-16;
Matches   50;  Conservative   0;  Mismatches   0;  Indels   0;  <

Qy   527 CCAGGGCCCTGCGCTACACACCGCCAGATCGGGAGTTTATAGTCACCCGG 576
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 CCAGGGCCCTGCGCTACACACCGCCAGATCGGGAGTTTATAGTCACCCGG 50

RESULT 29
AC129071/c
LOCUS       AC129071              163542 bp    DNA     linear   HTG 1
DEFINITION  Pan troglodytes clone RP43-149M23, WORKING DRAFT SEQUENCE
            ordered pieces.
ACCESSION   AC129071
VERSION     AC129071.2  GI:26449072
KEYWORDS    HTG; HTGS_PHASE2; HTGS DRAFT.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutel
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
            1 (bases 1 to 163542)
REFERENCE
  AUTHORS   Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S
            Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C.,
            Cariaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gu
            Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karl
            Laric,P., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,
            Margulies,E.H., Masiello,C., Maskeri,B., McDowell,J.,
            Paguirigan,C., Pearson,R., Portnoy,M.E., Prasad,A.,
            Reddix-Dugue,N., Schandler,K., Schueler,M.G., Sison,C.,
            Stantripop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., v
            Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
  TITLE     NISC Comparative Sequencing Initiative
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 163542)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUL-2002) NIH Intramural Sequencing Center,
            Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE   3 (bases 1 to 163542)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-DEC-2002) NIH Intramural Sequencing Center,
            Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT     On Dec 11, 2002 this sequence version replaced gi:2195501
            ------------ Genome Center
                Center: NIH Intramural Sequencing Center
                Center code: NISC
                Web site: http://www.nisc.nih.gov
```

```
             /note="assembly_fragment"
     24417..36979
             /note="assembly_fragment"
     37080..42619
             /note="assembly_fragment"
     42720..46464
             /note="assembly_fragment"
     46565..61551
             /note="assembly_fragment"
     61652..99225
             /note="assembly_fragment"
     99326..110459
             /note="assembly_fragment"
     110560..147587
             /note="assembly_fragment"
     147688..149736
             /note="assembly_fragment
             clone_end:SP6
             vector_side:right"

larity 100.0%;  Score 53;  DB 2;  Length 149736;
Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

CCAGGGTCCTCCCTGCGGATCCGCACCCTCCCCTGGGCCCATCTCAAGGC 812
||||||||||||||||||||||||||||||||||||||||||||||||||
CCAGGGTCCTCCCTGCGGATCCGCACCCTCCCCTGGGCCCATCTCAAGGC 43768

395                    50 bp     DNA     linear   PAT 30-AUG-2001
ce 74 from Patent WO0153486.
395.1  GI:15391196

etic construct
etic construct
cial sequences.

azi,A.J., Goddard,A., Godowski,P.J., Gurney,A.L.,
n,K.J., Marsters,S.A., Pan,J., Pitti,R.M., Roy,M.A., Smith,V.,
,D.M., Watanabe,C.K. and Wood,W.I.
sitions and methods for the treatment of tumour
=: WO 0153486-A 74 26-JUL-2001;
tech, Inc. (US)
             Location/Qualifiers
     1..50
             /organism="synthetic construct"
             /mol_type="unassigned DNA"
             /db_xref="taxon:32630"
             /note="Synthetic Oligonucleotide Probe."

larity 100.0%;  Score 50;  DB 6;  Pred. No. 7e-16;  Length 50;
Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

GCCCTCTGCGCTACAACCGCCAGATCGGGAGTTTATAGTCACCCGG 576
||||||||||||||||||||||||||||||||||||||||||||||
GCCCTCTGCGCTACAACCGCCAGATCGGGAGTTTATAGTCACCCGG 50

954
ligand polypeptide.    50 bp     DNA     linear   PAT 27-AUG-2002
954
954.1  GI:22636564
01522584-A/3.
ntified
ntified
```

ntact: nisc_zoo@nhgri.nih.gov
----- Project Information
nter project name: cmt
nter clone name: 149M23

quence data in this record represents an 'enhanced'
1 of a Phase 2 submission. Specifically, the indicated
and orientation of each sequence contig has been
ished using one or more of the following: read-pair
rom individual subclones, overlaps with neighboring
, alignment with available reference sequence (e.g.,
, and/or confirmation by PCR testing. In addition,
quence assembly is based on at least 8X average
ge in Q20 bases and has been reviewed to rule out
nisassemblies, the low-quality ends of sequence
s have been trimmed away, and each base is associated
Phrap-derived quality score.
------ Summary Statistics

sequencing vector: plasmid; n/a; 100% of reads
hemistry: Dye-terminator Big Dye; 100% of reads
sembly program: Phrap; version 0.990319
onsensus quality: 159675 bases at least Q40
onsensus quality: 160848 bases at least Q30
onsensus quality: 161608 bases at least Q20
nsert size: 185000; agarose-fp
nsert size: 162442; sum-of-contigs
ality coverage: 7.43x in Q20 bases; agarose-fp
ality coverage: 8.47x in Q20 bases; sum-of-contigs

: This is a 'working draft' sequence. It currently
ists of 12 contigs. Gaps between the contigs
represented as runs of N. The order of the pieces
elieved to be correct as given, however the sizes
he gaps between them are based on estimates that have
ided by the submittor.
sequence will be replaced
he finished sequence as soon as it is available and
accession number will be preserved.

```
1      10517: contig of 10517 bp in length
0518   10617: gap of unknown length
0618   25192: contig of 14575 bp in length
5193   25292: gap of unknown length
5293   27192: contig of 1900 bp in length
7193   27292: gap of unknown length
7293   29759: contig of 2467 bp in length
9760   29859: gap of unknown length
9860   57664: contig of 28005 bp in length
7865   57964: gap of unknown length
7965   61287: contig of 3323 bp in length
1288   61387: gap of unknown length
1388   84039: contig of 22652 bp in length
4040   84139: gap of unknown length
4140   114016: contig of 29877 bp in length
4017   114116: gap of unknown length
4117   116474: contig of 2358 bp in length
6475   116574: gap of unknown length
6575   121369: contig of 4795 bp in length
1370   121469: gap of unknown length
1470   154199: contig of 32730 bp in length
4200   154299: gap of unknown length
4300   163542: contig of 9243 bp in length.
```

                     Location/Qualifiers
                     1..163542
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /clone="RP43-149M23"
                     /clone_lib="RP43"
                     1..10517
                     /note="assembly_fragment
                     missing T7 clone end on 5' end of insert"
                     10618..25192
                     /note="assembly_fragment"

     misc_feature   25293..27192
                     /note="assembly_fragment"
     misc_feature   27293..29759
                     /note="assembly_fragment"
     misc_feature   29860..57864
                     /note="assembly_fragment"
     misc_feature   57965..61287
                     /note="assembly_fragment"
     misc_feature   61388..84039
                     /note="assembly_fragment"
     misc_feature   84140..114016
                     /note="assembly_fragment"
     misc_feature   114117..116474
                     /note="assembly_fragment"
     misc_feature   116575..121369
                     /note="assembly_fragment"
     misc_feature   121470..154199
                     /note="assembly_fragment"
     misc_feature   154300..163542
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:right"

ORIGIN

Query Match   3.5%;   Score 48;   DB 2;   Length 163542;
Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches   48; Conservative   0; Mismatches   0; Indels   0;   G

```
Qy    1312 CCTGTGGATTTTAAAACAGATATATTTTTTATTATTATTGTGACAAA 1359
           |||||||||||||||||||||||||||||||||||||||||||||||
Db  163542 CCTGTGGATTTTAAAACAGATATATTTTTTATTATTATTGTGACAAA 163495
```

RESULT 30
BD062757
LOCUS   BD062757   1168 bp   DNA   linear   PAT 27
DEFINITION A tumor necrosis factor related ligand.
ACCESSION   BD062757
VERSION   BD062757.1   GI:22608360
KEYWORDS   JP 2001505407-A/1.
SOURCE   unidentified
  ORGANISM unidentified
            unclassified.
REFERENCE   1 (bases 1 to 1168)
  AUTHORS   Chicheportiche,Y. and Browning,J.L.
  TITLE   A tumor necrosis factor related ligand
  JOURNAL   Patent: JP 2001505407-A 1 24-APR-2001;
            BIOGEN INC,THE FACULTY OF MEDICINE OF THE UNIVERSITY OF GE
COMMENT   OS TNF family related protein
            PN JP 2001505407-A/1
            PD 24-APR-2001
            PF 07-AUG-1997 JP 1998508239
            PR 07-AUG-1996 US 60/023541,18-OCT-1996 US   60/0285
            18-MAR-1997 US   60/040820
            PI YVES CHICHEPORTICHE,JEFFREY L BROWNING
            PC C12N15/28,C07K14/525,G01N33/68,C07K16/24,C12N15/11,A6
            PC C12N5/10,
            PC A61K39/395,A61K38/19,C07K14/705,C12N15/12
            CC Strandedness: Double;
            CC Topology: Linear;
            FH Key   Location/Qualifiers
            FT CDS   2..676.

FEATURES   Location/Qualifiers
     source   1..1168
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"

ORIGIN

Query Match   3.4%;   Score 46;   DB 6;   Length 1168;
Best Local Similarity 100.0%; Pred. No. 1.6e-13;
Matches   46; Conservative   0; Mismatches   0; Indels   0;   C

TGGGCTCTACTACCTGTACTGTCAGGTGCACTTTGATGAGGG 620
||||||||||||||||||||||||||||||||||||||||||
TGGGCTCTACTACCTGTACTGTCAGGTGCACTTTGATGAGGG 444

09   234801 bp   DNA   linear   HTG 13-NOV-2002
norvegicus clone CH230-212O18, *** SEQUENCING IN PROGRESS
unordered pieces.
09
09.6 GI:24941765
[TGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
norvegicus (Norway rat)
norvegicus
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
a.

ases 1 to 234801)
D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
ebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
n,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
lo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
;,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
nas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
o,J., Chavez,D., Chen,G., Chen,Y., Chen,Z., Chu,J.,
land,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
i,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
lo,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
;,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
;, Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
ndez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
c,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
eorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
atne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
ndez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
ns,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
on,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
chy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,J., Li,J., Li,Z., Liu,J.,
., Liu,Y., Liu,Y., London,P., Longacre,S., Lopez,J.,
suhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
hwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
m,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
y,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
ney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
avljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
rs,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
n,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
y,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
rvis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
elemeh,O., Okwuonu,G., Olanrpunsagoon,A., Pal,S., Parks,K.,
le,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
rnak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
r,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
er,F., Poindexter,C., Popovic,D., Primus,E., Pu,L.-L.,
R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Q., Wang,S., Warren,J., Warren,M., Wei,X., White,F.,
ams,G., Willson,R., Wieczyk,R., Wooden,H., Worley,K.,
t,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
, Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
rhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
stock,G. and Gibbs,R.A.
t Submission
lished
ases 1 to 234801)
y,K.C.
t Submission

JOURNAL   Submitted (15-APR-2002) Human Genome Sequencing Center, De
           of Molecular and Human Genetics, Baylor College of Medicir
           Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 234801)
AUTHORS   Rat Genome Sequencing Consortium.
TITLE   Direct Submission
JOURNAL   Submitted (13-NOV-2002) Human Genome Sequencing Center, De
           of Molecular and Human Genetics, Baylor College of Medicir
           Baylor Plaza, Houston, TX 77030, USA
COMMENT   On Nov 13, 2002 this sequence version replaced gi:23267374
           The sequence in this assembly is a combination of BAC base
           and whole genome shotgun sequencing reads assembled using
           (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig (
           in the feature table below represents a scaffold in the At
           assembly (a 'contig-scaffold'). Within each contig-scaffo
           individual sequence contigs are ordered and oriented, and
           by sized gaps filled with Ns to the estimated size. The se
           may extend beyond the ends of the clone and there may be r
           contigs within a contig-scaffold that consist entirely of
           genome shotgun sequence reads. Both end sequences and who
           shotgun sequence only contigs will be indicated in the fea
           table.
           --------------- Genome Center
             Center: Baylor College of Medicine
             Center code: BCM
             Web site: http://www.hgsc.bcm.tmc.edu/
             Contact: hgsc-help@bcm.tmc.edu
           --------------- Project Information
             Center project name: GUTA
             Center clone name: CH230-212O18
           --------------- Summary Statistics
             Assembly program: Phrap; version 0.990329
             Consensus quality: 227577 bases at least Q40
             Consensus quality: 230382 bases at least Q30
             Consensus quality: 231817 bases at least Q20
             Estimated insert size: 239036; sum-of-contigs estimat
             Quality coverage: 5x in Q20 bases; sum-of-contigs est
           * NOTE: Estimated insert size may differ from sequence le
           *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_d
           * NOTE: This is a 'working draft' sequence. It currently
           *   consists of 5 contigs. The true order of the pieces
           *   is not known and their order in this sequence record is
           *   arbitrary. Gaps between the contigs are represented as
           *   runs of N, but the exact sizes of the gaps are unknown.
           *   This record will be updated with the finished sequence
           *   as soon as it is available and the accession number wil
           *   be preserved.
           *   1   162694: contig of 162694 bp in length
           * 162695   162794: gap of unknown length
           * 162795   225625: contig of 62831 bp in length
           * 225626   225725: gap of unknown length
           * 225726   231241: contig of 5516 bp in length
           * 231242   231341: gap of unknown length
           * 231342   232640: contig of 1299 bp in length
           * 232641   232740: gap of unknown length
           * 232741   234801: contig of 2061 bp in length
FEATURES   Location/Qualifiers
     source   1..234801
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-212O18"
     misc_feature   complement(160419..161185)
                     /note="clone_boundary
                     clone_end:T7
                     site:
                     end sequence:BH264302"
     misc_feature   225726..227372
                     /note="wgs_end_extension
                     clone_end:T7"
ORIGIN

```
Qy    arity     2.3%;   Score 31;   DB 2;   Length 234801;
      100.0%;   Pred. No. 6e-05;
      onservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy    TGGGCTCTACTACCTGTACTGTCAGGT 605
      ||||||||||||||||||||||||||||
Db    TGGGCTCTACTACCTGTACTGTCAGGT 228959

      59                          26 bp    DNA       linear    PAT 30-AUG-2001
      ce 138 from Patent WO0153486.
      59    GI:15391274

      tic construct
      tic construct
      cial sequences.

REF-  azi,A.J., Goddard,A., Godowski,P.J., Gurney,A.L.,
      .K.J., Marsters,S.A., Pan,J., Pitti,R.M., Roy,M.A., Smith,V.,
      D.M., Watanabe,C.K. and Wood,W.I.
      itions and methods for the treatment of tumour
      : WO 0153486-A 138 26-JUL-2001
      ech, Inc. (US)
      Location/Qualifiers
      1..26
                /organism="synthetic construct"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32630"
                /note="Synthetic Oligonucleotide Probe."

      arity     1.9%;   Score 26;   DB 6;   Length 26;
      100.0%;   Pred. No. 0.017;
      onservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy    TCCCACATAAATACAGTATTCC 1030
      |||||||||||||||||||||
Db    TCCCACATAAATACAGTATTCC 1

      72                          140 bp   DNA       linear    PAT 24-OCT-2002
      ce 542 from Patent WO02064731.
      72    GI:24411826

      apiens (human)
      apiens
      ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

      an,A., Amson,R., Tuijnder,M. and Susini,L.
      ces involved in phenomena of tumour suppression, tumour
      ion, apoptosis and/or virus resistance and their use as
      neg
      : WO 02064731-A 542 22-AUG-2002;
      lar Engines Laboratories (FR)
      Location/Qualifiers
      1..140
                /organism="Homo sapiens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:9606"

      arity     1.9%;   Score 26;   DB 6;   Length 140;
      100.0%;   Pred. No. 0.02;
      onservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy    CATTGTTAGACTTTGAAATTC 26
```

```
Db    |||||||||||||||||||||||||| ATGTCATTGTTAGACTTTGAAATTC 42
      67

RESULT 34
AX522345
LOCUS       AX522345                        145 bp   DNA       linear    PAT 24
DEFINITION  Sequence 15 from Patent WO02064731.
ACCESSION   AX522345
VERSION     AX522345.1   GI:24411299
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1
REFERENCE
  AUTHORS   Telerman,A., Amson,R., Tuijnder,M. and Susini,L.
  TITLE     Sequences involved in phenomena of tumour suppression, tum
            reversion, apoptosis and/or virus resistance and their use
            medicines
  JOURNAL   Patent: WO 02064731-A 15 22-AUG-2002;
            Molecular Engines Laboratories (FR)
FEATURES    Location/Qualifiers
  source    1..145
                /organism="Homo sapiens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:9606"
ORIGIN

Query Match              1.9%;   Score 26;   DB 6;   Length 145;
Best Local Similarity   100.0%;   Pred. No. 0.02;
Matches     26;   Conservative   0;   Mismatches   0;   Indels   0;   G

Qy    1  ATGTCATTGTTAGACTTTGAAATTC 26
         ||||||||||||||||||||||||||
Db   74  ATGTCATTGTTAGACTTTGAAATTC 99

RESULT 35
AX522465
LOCUS       AX522465                        145 bp   DNA       linear    PAT 24
DEFINITION  Sequence 135 from Patent WO02064731.
ACCESSION   AX522465
VERSION     AX522465.1   GI:24411419
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1
REFERENCE
  AUTHORS   Telerman,A., Amson,R., Tuijnder,M. and Susini,L.
  TITLE     Sequences involved in phenomena of tumour suppression, tum
            reversion, apoptosis and/or virus resistance and their use
            medicines
  JOURNAL   Patent: WO 02064731-A 135 22-AUG-2002;
            Molecular Engines Laboratories (FR)
FEATURES    Location/Qualifiers
  source    1..145
                /organism="Homo sapiens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:9606"
ORIGIN

Query Match              1.9%;   Score 26;   DB 6;   Length 145;
Best Local Similarity   100.0%;   Pred. No. 0.02;
Matches     26;   Conservative   0;   Mismatches   0;   Indels   0;   C

Qy    1  ATGTCATTGTTAGACTTTGAAATTC 26
         ||||||||||||||||||||||||||
Db   74  ATGTCATTGTTAGACTTTGAAATTC 99

RESULT 36
```

ORGANIZM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 353)
AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
           Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schu
           Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bha
           Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsie
           Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,
           Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
           Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
           Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters
           Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
           McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
           Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
           Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R
           Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
           Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y
           Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
           Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Mye
           Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.I
           Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE   Generation and initial analysis of more than 15,000 full-
           human and mouse cDNA sequences
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE   22388257
PUBMED   12477932
REFERENCE 2 (bases 1 to 353)
AUTHORS   Strausberg,R.
TITLE   Direct Submission
JOURNAL   Submitted (17-NOV-2003) National Institutes of Health, Ma
           Gene Collection (MGC), Cancer Genomics Office, National C
           Institute, 31 Center Drive, Room 11A03, Bethesda, MD 2089
           USA
REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: ATCC/DCTD/DTP
           cDNA Library Preparation: CLONTECH Laboratories, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Genome Sequence Centre,
           BC Cancer Agency, Vancouver, BC, Canada
           info@bcgsc.bc.ca
           Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfie
           Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ra
           Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver L
           Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy,
           Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, J
           Schein, Duane Smailus, Michael Smith, Lorraine Spence, Je
           Michael Thorne, Miranda Tsai, Natasja van den Bosch, Jil
           George Yang, Scott Zuyderduyn, Marco Marra.

           Clone distribution: MGC clone distribution information ca
           through the I.M.A.G.E. Consortium/LLNL at: http://image.l
           Series: IRAL Plate: 51 Row: c Column: 23
           This clone was selected for full length sequencing becaus
           passed the following selection criteria: Hexamer frequenc
           analysis.
FEATURES   Location/Qualifiers
     source   1..353
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3932215"
                     /tissue_type="Skin, melanotic melanoma, high MDR
                     /clone_lib="NIH_MGC_62"
                     /lab_host="DH10B"
                     /note="Vector: pDNR-LIB"
ORIGIN

Query Match   1.9%; Score 26; DB 9; Length 353;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches   26; Conservative   0; Mismatches   0; Indels   0;

---

95 168 from Patent EP1162276.
95   153 bp   DNA   linear   PAT 07-JAN-2002
95.1 GI:18093362

sapiens (human)
sapiens
yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
lia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

,M.D., Winter,H. and Reinartz,J.
olecules to be used as indicators of the functional and
ational state of t-lymphocytes
:: EP 1162276-A 168 12-DEC-2001;
herapeutics GmbH (DE)
          Location/Qualifiers
     1..153
                /organism="Homo sapiens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:9606"

larity   1.9%; Score 26; DB 6; Length 153;
Conservative   0; Mismatches   0; Indels   0; Gaps   0;

TCATTGTTAGACTTTGAAATTTC 26
|||||||||||||||||||||||
TCATTGTTAGACTTTGAAATTTC 55

---

050 890 from Patent WO0157207.
nce   281 bp   DNA   linear   PAT 31-AUG-2001
050.1 GI:15423473

sapiens (human)
sapiens
yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
lia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

e,P.A. and Mannion,J.
sitions and methods for the therapy and diagnosis of ovarian
r
t: WO 0157207-A 890 09-AUG-2001;
A CORPORATION (US)
          Location/Qualifiers
     1..281
                /organism="Homo sapiens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:9606"

larity   1.9%; Score 26; DB 6; Length 281;
Conservative   0; Mismatches   0; Indels   0; Gaps   0;

TCATTGTTAGACTTTGAAATTTC 26
|||||||||||||||||||||||
TCATTGTTAGACTTTGAAATTTC 245

---

447   353 bp   mRNA   linear   PRI 18-NOV-2003
sapiens cDNA clone IMAGE:3932215, partial cds.
447
447.1 GI:38383152

sapiens (human)

Query Match   1.9%; Score 26; DB 6; Length 387;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 26; Conservative 0; Mismatches 0; Indels 0; G:

Qy   1 ATGTCATTGTTAGACTTTGAAATTTC 26
          ||||||||||||||||||||||||||
Db   331 ATGTCATTGTTAGACTTTGAAATTTC 356

RESULT 41
AX198506
LOCUS   AX198506   391 bp   DNA   linear   PAT 29
DEFINITION   Sequence 961 from Patent WO0151513.
ACCESSION   AX198506
VERSION   AX198506.1 GI:15388827
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Algate,P.A.
  TITLE   Ovarian tumor-associated sequences
  JOURNAL   Patent: WO 0151513-A 961 19-JUL-2001;
             CORIXA CORPORATION (US)
FEATURES   Location/Qualifiers
     source   1..391
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match   1.9%; Score 26; DB 6; Length 391;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 26; Conservative 0; Mismatches 0; Indels 0; G

Qy   1 ATGTCATTGTTAGACTTTGAAATTTC 26
          ||||||||||||||||||||||||||
Db   220 ATGTCATTGTTAGACTTTGAAATTTC 245

RESULT 42
AR427022
LOCUS   AR427022   416 bp   DNA   linear   PAT 18
DEFINITION   Sequence 18519 from patent US 6639063.
ACCESSION   AR427022
VERSION   AR427022.1 GI:40182132
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE   1 (bases 1 to 416)
  AUTHORS   Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
  TITLE   EST's and encoded human proteins
  JOURNAL   Patent: US 6639063-A 18519 28-OCT-2003;
FEATURES   Location/Qualifiers
     source   1..416
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match   1.9%; Score 26; DB 6; Length 416;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 26; Conservative 0; Mismatches 0; Indels 0; G

Qy   1 ATGTCATTGTTAGACTTTGAAATTTC 26
          ||||||||||||||||||||||||||
Db   331 ATGTCATTGTTAGACTTTGAAATTTC 356

RESULT 43
BD122575
LOCUS   BD122575   416 bp   DNA   linear   PAT 18

ATTGTTAGACTTTGAAATTTC 26
|||||||||||||||||||||
ATTGTTAGACTTTGAAATTTC 291

                      387 bp   DNA   linear   PAT 18-DEC-2003

e 18517 from patent US 6639063.

0.1 GI:40182130

ified.
,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
nd encoded human proteins
US 6639063-A 18517 28-OCT-2003;
             Location/Qualifiers
             1..387
                     /organism="unknown"
                     /mol_type="genomic DNA"

arity 100.0%; Pred. No. 0.022;
nservative 0; Mismatches 0; Indels 0; Gaps 0;

CATTGTTAGACTTTGAAATTTC 26
||||||||||||||||||||||
CATTGTTAGACTTTGAAATTTC 356

                      387 bp   DNA   linear   PAT 18-SEP-2002

 encoded human protein.

73.1 GI:23217518
010789-A/14650.
apiens (human)
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 387)
,J.B.D.M., Jobert,S. and Giordano,J.E.
 encoded human protein
 JP 2002010789-A 14650 15-JAN-2002;
CORP

omo sapiens (human)
 2002010789-A/14650
5-JAN-2002
7-AUG-2000 JP 2000280989
5-AUG-1999 US   60/147499
EAN BAPUTIST DUMAS MILNE EDWARDS,SEVELIN JOBERT,JEAN EVE PI
ANO
2N15/09,C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19, PC
/21,
2N5/10,C12P21/02,C12P21/08,C12Q1/68,C12N15/00,C12N5/00, PC
/00
ST and encoded human protein
ey   1..387
ource   Location/Qualifiers
                /organism='Homo sapiens (human)'.

             Location/Qualifiers
             1..387
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

d encoded human protein.
75 GI:23217520
75.1 GI:23217520
2010789-A/14652.
apiens (human)
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 416)
s,J.B.D.M.,Jobert,S. and Giordano,J.E.
d encoded human protein
: JP 2002010789-A 14652 15-JAN-2002;
: CORP

Homo sapiens (human)
P 2002010789-A/14652
5-JAN-2002
07-AUG-2000 JP 2000280989
05-AUG-1999 US 60/147499
JEAN BAPUTIST DUMAS MILNE EDWARDS,SEVELIN JOBERT,JEAN EVE PI
DANO
C12N15/09,C12N15/09,C07K16/18,C12N1/15,C12N1/19, PC
1/21,
C12N5/10,C12P21/02,C12P21/08,C12Q1/68,C12N15/00,C12N5/00, PC
5/00
EST and encoded human protein
Key   Location/Qualifiers
source   1..416
         /organism='Homo sapiens (human)'.

         Location/Qualifiers
         1..416
         /organism="Homo sapiens"
         /mol_type="genomic DNA"
         /db_xref="taxon:9606"

larity   1.9%;   Score 26;   DB 6;   Length 416;
Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

TCATTGTTAGACTTTGAAATTTC 26
||||||||||||||||||||||||
TCATTGTTAGACTTTGAAATTTC 356

480   418 bp   mRNA   linear   PRI 27-JUN-2002
sapiens, clone IMAGE:5214272, mRNA, partial cds.
480
480.1 GI:21619102

sapiens (human)
sapiens
yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
lia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ases 1 to 418)
sberg,R.
t Submission
tted (06-JUN-2002) National Institutes of Health, Mammalian
Collection (MGC), Cancer Genomics Office, National Cancer
tute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

MGC Project URL: http://mgc.nci.nih.gov
act: MGC help desk
: cgapbs-r@mail.nih.gov
e Procurement: Life Technologies, Inc.
Library Preparation: Life Technologies, Inc.
Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Sequencing by: National Institutes of Health Intramural
ancing Center (NISC),
hersburg, Maryland;
site: http://www.nisc.nih.gov/
act: nisc_mgc@nhgri.nih.gov

Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Broc
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legas
Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McClos
McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touch
Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggir
Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information car
through the I.M.A.G.E. Consortium/LLNL at: http://image.ll
Series: IRAK Plate: 64 Row: 1 Column: 24
This clone was selected for full length sequencing because
passed the following selection criteria: Hexamer frequenc)
analysis.

FEATURES   Location/Qualifiers
source   1..418
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:5214272"
             /tissue_type="Blood, adult leukocytes"
             /clone_lib="NIH_MGC_118"
             /lab_host="DH10B"
             /note="Vector: pCMV-SPORT6"
CDS   <1..281
             /codon_start=3
             /product="Unknown (protein for IMAGE:5214272)"
             /protein_id="AAH32480.1"
             /db_xref="GI:21619103"
             /translation="RSVVLLVAVRLHTLLSCPLEQPAGTEWILEEGV
             DIYNLRSPDENWRWNMRGALWKEKDRPCAFMKVKIWLNQFHKVTVYIA

ORIGIN

Query Match   1.9%;   Score 26;   DB 9;   Length 418;
Best Local Similarity 100.0%; Pred. No. 0.022; Indels   0;
Matches 26; Conservative   0; Mismatches   0; Indels   0;

Qy   1 ATGTCATTGTTAGACTTTGAAATTTC 26
         ||||||||||||||||||||||||||
Db   309 ATGTCATTGTTAGACTTTGAAATTTC 334

RESULT 45
AX330518/c
LOCUS   AX330518   439 bp   DNA   linear   PAT 1
DEFINITION   Sequence 1027 from Patent WO0194629.
ACCESSION   AX330518
VERSION   AX330518.1 GI:18103496
KEYWORDS
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutel
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
REFERENCE   1
AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,(
             Horrigan,S., Soppet,D.R. and Weaver,Z.
TITLE   Cancer gene determination and therapeutic screening usinc
             gene sets
JOURNAL   Patent: WO 0194629-A 1027 13-DEC-2001;
             Avalon Pharmaceuticals (US)
FEATURES   Location/Qualifiers
source   1..439
             /organism="Homo sapiens"
             /mol_type="unassigned DNA"
             /db_xref="taxon:9606"

ORIGIN

Query Match   1.9%;   Score 26;   DB 6;   Length 439;
Best Local Similarity 100.0%; Pred. No. 0.022; Indels   0;
Matches 26; Conservative   0; Mismatches   0; Indels   0;

Qy   1 ATGTCATTGTTAGACTTTGAAATTTC 26

ACCESSION   BD030667
VERSION   BD030667.1 GI:22572409
KEYWORDS   JP 2001269182-A/6913 .
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 452)
  AUTHORS   Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
  TITLE   Sequence tag and encoded human protein
  JOURNAL   Patent: JP 2001269182-A 6913 02-OCT-2001;
    GENSET
COMMENT   OS   Homo sapiens (human)
    PN   JP 2001269182-A/6913
    PD   02-OCT-2001
    PF   24-FEB-2000 JP 2000118773
    PR   26-FEB-1999 US   60/122487
    PI   JEAN BAPTIST DUMAS MILNE EDWARDS, EIMERIC DUCLAIR,JEA
    PI   JORDAN
    PC   C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N
    C12N5/10,
    PC   C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12
    G06F15/40
    CC
FH   Key   Location/Qualifiers.
FEATURES   Location/Qualifiers
    source   1..452
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
ORIGIN

Query Match   1.9%;   Score 26;   DB 6;   Length 452;
Best Local Similarity   100.0%;   Pred. No. 0.022;
Matches   26;   Conservative   0;   Mismatches   0;   Indels   0;   C

Qy   1 ATGTCATTGTTAGACTTTGAAATTTC 26
      ||||||||||||||||||||||||||
Db   414 ATGTCATTGTTAGACTTTGAAATTTC 439

RESULT 49
AX381620
LOCUS   AX381620   483 bp   DNA   linear   PAT 18
DEFINITION   Sequence 558 from Patent WO0212280.
ACCESSION   AX381620
VERSION   AX381620.1 GI:19576442
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
REFERENCE   1
  AUTHORS   Pyle,R.A., Xu,J. and Secrist,H.
  TITLE   Compositions and methods for the therapy and diagnosis of
    cancer
  JOURNAL   Patent: WO 0212280-A 558 14-FEB-2002;
    CORIXA CORPORATION (US)
FEATURES   Location/Qualifiers
    source   1..483
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"
ORIGIN

Query Match   1.9%;   Score 26;   DB 6;   Length 483;
Best Local Similarity   100.0%;   Pred. No. 0.022;
Matches   26;   Conservative   0;   Mismatches   0;   Indels   0;

Qy   1 ATGTCATTGTTAGACTTTGAAATTTC 26
      ||||||||||||||||||||||||||
Db   397 ATGTCATTGTTAGACTTTGAAATTTC 422

|||||||||||||||||||||||||||
ATTGTTAGACTTTGAAATTTC 47

    439 bp   DNA   linear   PAT 14-JUN-2002
7   e 2744 from Patent WO0229103.
7
7.1 GI:21442802

apiens (human)
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
pression profiles in liver cancer
: WO 0229103-A 2744 11-APR-2002;
OGIC INC (US)
    Location/Qualifiers
    1..439
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"
    /note="EMBL/GenBank Accession No. N98464"

arity   100.0%;   Pred. No. 0.022;   Indels   0;   Gaps   0;
onservative   0;   Mismatches   0;

CATTGTTAGACTTTGAAATTTC 26
||||||||||||||||||||||
CATTGTTAGACTTTGAAATTTC 47

34   452 bp   DNA   linear   PAT 18-DEC-2003
ce 10997 from Patent EP1033401.
34
34.1 GI:40050018

apiens (human)
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
sed sequence tags and encoded human proteins
: EP 1033401-A 10997 06-SEP-2000;
 (FR)
    Location/Qualifiers
    1..452
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"

arity   100.0%;   Pred. No. 0.022;   Indels   0;   Gaps   0;
onservative   0;   Mismatches   0;

TCATTGTTAGACTTTGAAATTTC 26
|||||||||||||||||||||||
TCATTGTTAGACTTTGAAATTTC 439

    452 bp   DNA   linear   PAT 27-AUG-2002
667
nce tag and encoded human protein.

...dard; DNA; 153 BP.

(first entry)

...ell derived DNA fragment #168.

...ell; immunosuppressive; immunostimulant; antinflammatory;
...ene therapy; vaccine; allergen; transplant rejection;
...host disease; malignant disease; ds.

.

2000DE-01021834.

2000DE-01021834.

THERAPEUTICS GMBH.

...inter H, Reinartz J;

320/04.

...cative of T cell activation and functional status, useful
...and therapy e.g. of autoimmunity or transplant rejection.

...48; 94pp; German.

...represents a novel messenger RNA, (mRNA), (I), for use as
...the activation and functional status of T cells, that have
...reduced expression, and are present at higher or lower
..., in activated T cells, relative to normal or resting cells,
...ridizes to any of 334 sequences, reproduced, or their
...complements or fragments. The products of the invention have
...sive, immunostimulant, antinflammatory and cytostatic
...can be used for gene therapy. The polynucleotides of the
...used: (i) as reagent for detecting activation/functional
...ells, for diagnosis, therapy, modulation or control of the
...ses of (auto)immunity (against microorganisms, vaccines or
... transplant rejection; immunologically-related inflammation;
...sion; immune deficiency; guest versus host disease, and
...eases of the immune system; (ii) for identifying agents,
...rmaceuticals, that bind to (II) or identifying polypeptides
...to prepare kits for measuring gene expression profiles in
...ne, especially T, cells; (iv) to raise antibodies (Ab)
...nst (III); and (v) to prepare binding molecules (IV)
...(II). Ab and (IV) are also useful for detecting and
...e activation and functional status of T cells. AAI68865-
...esent the activated T-cell derived polynucleotide fragments
...the method of the invention

...BP; 43 A; 23 C; 28 G; 59 T; 0 U; 0 Other;

...arity   100.0%;   Pred. No. 0.11;   Indels   0;   Gaps   0;
...onservative   0;   Mismatches   0;

...TCATTGTTAGACTTTGAAATTTC 26
...|||||||||||||||||||||||
...TCATTGTTAGACTTTGAAATTTC 55

...ndard; cDNA; 281 BP.

DT   07-NOV-2001 (first entry)
XX
DE   Human ovarian PCR-subtracted cDNA library clone #890.
XX
KW   Immunogenic protein; cancer; ovarian tumour; T-cell stimulation;
KW   gene therapy; cytostatic; T-cell expansion; nucleic acid hybridis
KW   primer; probe.
XX
OS   Homo sapiens.
XX
PN   WO200157207-A2.
XX
PD   09-AUG-2001.
XX
PF   05-FEB-2001; 2001WO-US003733
XX
PR   04-FEB-2000; 2000US-0180403P.
PR   28-MAR-2000; 2000US-0192745P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Algate PA, Mannion J;
XX
DR   WPI; 2001-488879/53.
XX
PT   New polynucleotides encoding ovarian tumor proteins, useful for t
PT   ovarian cancer, and as probes, primers, and markers of cancer
PT   progression.
XX
PS   Example 1; Page 253; 378pp; English.
XX
CC   The invention comprises compositions used for the therapy and dia
CC   of ovarian cancer. The compositions comprise one or more ovarian
CC   proteins, their associated polynucleotides, or immunogenic portic
CC   the proteins. The ovarian tumour polynucleotides and polypeptides
CC   useful for stimulating and/or expanding T cells specific for a tu
CC   protein. They are also useful for inhibiting the development of c
CC   cells allowing them to proliferate, and administering to the pat:
CC   a patient with an ovarian tumour DNA or protein by incubating iso
CC   sequences can be used as markers for cancer, for example, to moni
CC   ovarian cancer progression. Probes and primers are useful in nucl
CC   hybridisation, in detecting the presence of complementary sequenc
CC   given sample, for preparing mutant species and for preparing othe
CC   genetic constructions. Sequences AAS23820-AAS25231 and AAS25328-4
CC   represent human ovarian tumour protein cDNA clones
XX
SQ   Sequence 281 BP; 98 A; 49 C; 59 G; 74 T; 0 U; 1 Other;

Query Match   1.9%;   Score 26;   DB 4;   Length 281;
Best Local Similarity   100.0%;   Pred. No. 0.11;
Matches   26;   Conservative   0;   Mismatches   0;   Indels   0;   (

Qy   1 ATGTCATTGTTAGACTTTGAAATTTC 26
         ||||||||||||||||||||||||||
Db   220 ATGTCATTGTTAGACTTTGAAATTTC 245

RESULT 40
ABQ60530
ID   ABQ60530 standard; cDNA; 386 BP.
XX
AC   ABQ60530;
XX
DT   02-AUG-2002 (first entry)
XX
DE   Human colon cancer related nucleotide sequence SEQ ID NO:4225.
XX
KW   Human; colon cancer; cancer; tissue profiling; forensic; mapping
KW   genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200229086-A2.

PI Algate PA;
XX
DR WPI; 2001-425866/45.
XX
PT Novel ovarian tumor proteins, and nucleic acids encoding them, use
PT treat and diagnose cancers, particularly ovarian cancer.
XX
PS Claim 5; Page 239; 338pp; English.
XX
CC AAH82377 to AAH83878 represent human ovarian tumour-associated
CC polynucleotide sequences which encode ovarian tumour proteins. The
CC ovarian tumour protein and polynucleotide sequences have cytostat:
CC activity, and can be used in gene therapy and vaccine production.
CC ovarian tumour proteins and polynucleotides can be used to inhibit
CC development of cancer, particularly ovarian cancer. They can also
CC to diagnose the onset and progression of cancer
XX
SQ Sequence 391 BP; 118 A; 80 C; 88 G; 97 T; 0 U; 8 Other;

Query Match 1.9%; Score 26; DB 5; Length 391;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 26; Conservative 0; Mismatches 0; Indels 0; G.

Qy 1 ATGTCATTGTTAGACTTTGAAATTTC 26
    ||||||||||||||||||||||||||
Db 220 ATGTCATTGTTAGACTTTGAAATTTC 245

RESULT 42
ABX74646/c
ID ABX74646 standard; cDNA; 425 BP.
XX
AC ABX74646;
XX
DT 21-MAR-2003 (first entry)
XX
DE Human cDNA sequence #113 up-regulated in CC-RCC patients.
XX
KW Human; microarray; solid surface; immobilised probe; CC-RCC;
KW differential expression profile; aggressive CC-RCC tumour type;
KW non-aggressive CC-RCC tumour type; clear cell renal carcinoma;
KW gene expression profiling; tumour tissue; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200279411-A2.
XX
PD 10-OCT-2002.
XX
PF 29-MAR-2002; 2002WO-US009576.
XX
PR 29-MAR-2001; 2001US-0279411P.
XX
PA (VAND-) VAN ANDEL INST.
XX
PI Haab B, Rhodes D, Teh BT, Takashi M;
XX
DR WPI; 2003-040679/03.
XX
PT New microarray, comprising a matrix of cDNA probe from a set of p
PT immobilized to a solid surface in predetermined order, useful in
PT prognosis of patients with clear cell renal carcinoma.
XX
PS Claim 35; SEQ ID NO 223; 179pp; English.
XX
CC The present invention relates to a microarray comprising a matrix
CC least one cDNA probe from a set of probes immobilised to a solid
CC in a predetermined order, where a row of pixels corresponds to re
CC of one distinct probe from the set. The probes are complementary
CC nucleic acid sequences that are expressed differentially in aggre
CC compared to non-aggressive types of clear cell renal carcinoma (C
CC and that hybridise to the probes under high stringency conditions
CC microarray is useful for the prognosis of patients with CC-RCC, w

---

001WO-US030732.
000US-0237271P.
CORP.
stle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
 Lewis ME;
15/45.
nucleic acid that is differentially expressed in cancer
 for determining the presence of colon cancer in a cell or
nd in antisense therapy.
-; 796pp; English.
BQ60787 represent isolated nucleic acids (I) differentially
cancer tissues. ABB78993 to ABB79004 represent proteins
a ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
nse therapy. An antibody immunoreactive with a polypeptide
is useful for detecting cancer in a patient sample, and
the presence or absence of a polynucleotide encoded by a
which hybridies to (I) in a cell. A probe/primer derived
be used for determining the presence of a nucleic acid which
(I), and for determining the phenotype of cells in a sample
a patient. (I) is useful for determining the presence of
n a cell or tissue type, for determining the presence or
 type of cancer, in antisense therapy, to generate
 a solid surface, to identify a chromosome on which the
gene resides, and in tissue profiling, forensics, genetic
ping and diagnostic application. (I) can be used to raise
nd to screen for peptide analogues and antagonists
BP; 128 A; 72 C; 85 G; 99 T; 0 U; 2 Other;
rity 1.9%; Score 26; DB 6; Length 386;
100.0%; Pred. No. 0.11;
onservative 0; Mismatches 0; Indels 0; Gaps 0;
CATTGTTAGACTTTGAAATTTC 26
CATTGTTAGACTTTGAAATTTC 343
dard; cDNA; 391 BP.
(first entry)
tumour associated polynucleotide sequence SEQ ID NO:961.
n tumour; ovarian cancer; diagnosis; gene therapy;
vaccine; ss.
2.
2001WO-US001575.
2000US-0176722P.
A CORP.

...d non-aggressive CC-RCC tumour types are characterised by
expression profiles of genes that hybridise with one or more
...lised on the microarray. The arrays are useful for gene
...ofiling of tumour and normal tissues. The present sequence
human cDNA sequence up-regulated in CC-RCC patients

BP; 118 A; 91 C; 78 G; 138 T; 0 U; 0 Other;

...tarity   1.9%;   Score 26; DB 7; Length 425;
...onservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy ...CATTGTTAGACTTTGAAATTTC 26
    ...||||||||||||||||||||||
Db ...CATTGTTAGACTTTGAAATTTC 45

...dard; DNA; 439 BP.

(first entry)

...arcinoma related gene sequence SEQ ID NO:1027.

...r; colon; breast; ovary; oesophagus; kidney; thyroid;
...g; prostate; pancreas; carcinoma; antitumour; cancerous;
...gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;

...A2.

2001WO-US010838.

2000US-0209473P.
2000US-0209531P.
2000US-0233133P.
2000US-0233617P.
2000US-0234009P.
2000US-0234034P.
2000US-0234052P.
2000US-0234509P.
2000US-0234567P.
2000US-0234923P.
2000US-0234924P.
2000US-0235077P.
2000US-0235082P.
2000US-0235134P.
2000US-0235280P.
2000US-0235637P.
2000US-0235638P.
2000US-0235711P.
2000US-0235720P.
2000US-0235840P.
2000US-0235863P.
2000US-0236028P.
2000US-0236032P.
2000US-0236033P.
2000US-0236034P.
2000US-0236109P.
2000US-0236111P.
2000US-0236842P.
2000US-0236891P.
2000US-0237172P.
2000US-0237173P.
2000US-0237278P.
2000US-0237294P.

PR 02-OCT-2000; 2000US-0237729SP.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrig.
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT agent to be tested for anti-neoplastic activity, and determining
PT in expression of a signature gene set.
XX
PS Claim 1; SEQ ID NO 1027; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for
CC neoplastic agent. The method involves exposing cells to a chemic.
CC to be tested for anti-neoplastic activity, determining a change
CC expression of at least one gene (I) of a signature gene set, whe
CC comprises a sequence (S) selected from 8447 sequences (given in
CC to ABL70110), or is at least 95% identical to (S), where a chang
CC expression is indicative of anti-neoplastic activity. (I) has cy
CC activity and can be used in gene therapy. M1 can be used for scr
CC anti-neoplastic agent, and can be used for producing a product w
CC the data collected with respect to the anti-neoplastic agent as
CC of M1, and the data is sufficient to convey the chemical structu
CC properties of the agent. M1 can be used in the treatment of canc
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer,
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma an
CC tumour
XX
SQ Sequence 439 BP; 118 A; 101 C; 83 G; 137 T; 0 U; 0 Other;

Query Match   1.9%; Score 26; DB 6; Length 439;
Best Local Similarity   100.0%; Pred. No. 0.11;
Matches 26; Conservative   0; Mismatches   0; Indels   0;

Qy   1 ATGTCATTGTTAGACTTTGAAATTTC 26
       ||||||||||||||||||||||||||
Db 72 ATGTCATTGTTAGACTTTGAAATTTC 47

RESULT 44
ABN96246/C
ID ABN96246 standard; DNA; 439 BP.
XX
AC ABN96246;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #2744 used to diagnose liver cancer.
XX
KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytostatic; expression profile; disease
KW disease progression; drug toxicity; drug efficacy; drug metaboli
XX
OS Homo sapiens.
XX
PN WO200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US030589.

PT diagnostic, forensic, gene therapy and chromosome mapping procedu
XX
PS Claim 1; SEQ ID NO 10997; 71pp + Sequence Listing; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived
CC mRNAs encoding secreted proteins. No ORF has yet been conclusivel
CC identified within the present sequence. The 5' ESTs were prepared
CC total human RNAs or polyA+ RNAs derived from 30 different tissues
CC sequences usually correspond mainly to the 3' untranslated region
CC of the mRNA because they are often obtained from oligo-dT primed
CC libraries. Such ESTs are not well suited for isolating cDNA seque
CC derived from the 5' ends of mRNAs and even in those cases where l
CC cDNA sequences have been obtained, the full 5' UTR is rarely incl
CC ESTs are derived from mRNAs with intact 5' ends and can therefore
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also us
CC diagnostic, forensic, gene therapy and chromosome mapping procedu
CC They are used to obtain upstream regulatory sequences and to desi
CC expression and secretion vectors
XX
SQ Sequence 452 BP; 122 A; 95 C; 112 G; 122 T; 0 U; 1 Other;

Query Match   1.9%; Score 26; DB 3; Length 452;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 26; Conservative 0; Mismatches 0; Indels 0; G

Qy   1 ATGTCATTGTTAGACTTGAAATTTC 26
         |||||||||||||||||||||||||
Db   414 ATGTCATTGTTAGACTTGAAATTTC 439

RESULT 46
ABV86720
ID ABV86720 standard; cDNA; 469 BP.
XX
AC ABV86720;
XX
DT 13-DEC-2002 (first entry)
XX
DE Human colon cancer related cDNA SEQ ID NO 31.
XX
KW Human; colon; cancer; cytostatic; tumour; gene therapy; vaccine;
KW ss.
XX
OS Homo sapiens.
XX
PN WO200258534-A2.
XX
PD 01-AUG-2002.
XX
PF 16-NOV-2001; 2001WO-US043704.
XX
PR 20-NOV-2000; 2000US-0252222P.
PR 06-FEB-2001; 2001US-0267011P.
PR 28-MAR-2001; 2001US-0279670P.
PR 10-JUL-2001; 2001US-0304037P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Stolk JA, Xu J, Chenault RA, Meagher MJ, Secrist H, King GE;
XX
DR WPI; 2002-608400/65.
XX
PT New isolated tumor colon polynucleotide and polypeptide, useful f
PT diagnosis, prevention and/or treatment of cancer, in particular c
PT cancer.
XX
PS Claim 1; SEQ ID NO 31; 266pp + Sequence Listing; English.
XX
CC The invention relates to a human colon tumour expressed polynucle
CC (I) encoding a polypeptide (II, ABP67991-ABP67996) comprising: (i
CC 2600 fully defined nucleotide sequences (ABV8669-ABV89289); (ii)
CC complements of (i); (iii) at least 20 contiguous residues of (i);
CC sequences that hybridize to (i), under moderately stringent condi

---

2000US-0237054P.
XX
LOGIC INC.
XX
ares C, Peres-Da-Silva S, Vockley JG;
XX
119/45.
XX
d detecting the progression of liver cancer, hepatocellular
metastatic liver tumor in a patient, involves detecting the
ession of two or more genes in a liver tissue sample.
XX
ID NO 2744; 298pp; English.
XX
relates to a novel method for diagnosing and detecting the
f liver cancer, hepatocellular carcinoma or metastatic liver
atient, and differentiating metastatic liver cancer from
r carcinoma in a patient, involving detecting the level of
two or more genes represented in ABN93503-ABN97455 in a
. The method of the invention has hepatotropic, and
ivity. The method is useful for diagnosing and detecting
on of liver cancer, hepatocellular carcinoma and metastatic
na in a patient. The method is useful for identifying
ofiles which serve as useful diagnostic markers as well as
can be used to monitor disease states, disease progression,
, drug efficacy and drug metabolism. Note: The sequence data
at did not form part of the printed specification, but was
lectronic format directly from WIPO at
pub/published_pct_sequences

BP; 118 A; 101 C; 83 G; 137 T; 0 U; 0 Other;

arity 1.9%; Score 26; DB 6; Length 439;
arity 100.0%; Pred. No. 0.11;
onservative 0; Mismatches 0; Indels 0; Gaps 0;

CATTGTTAGACTTGAAATTTC 26
||||||||||||||||||||||
CATTGTTAGACTTGAAATTTC 47

dard; cDNA; 452 BP.

(first entry)

d protein 5' EST, SEQ ID NO: 10997.

; expressed sequence tag; secreted protein; cDNA isolation;
chromosome mapping; ss.

2000EP-00200610.

99US-0122487P.

r.

dwards J, Duclert A, Giordano J;

381/45.

cid that is a 5' expressed sequence tag (5' EST) for
As and genomic DNAs that correspond to 5'ESTs and for

s having at least 75% or 90% identity to (i); or (vi)
ariants of (i). The compositions and methods of the present
e useful for the diagnosis, prevention and/or treatment of
icularly colon cancer. (I) can be used in gene therapy and
are useful in pharmaceutical compositions such as vaccines.
quence data for this patent did not form part of the printed
n, but was obtained in electronic format directly from WIPO
int/pub/published_pct_sequences

BP; 142 A; 94 C; 111 G; 122 T; 0 U; 0 Other;

larity   1.9%; Score 26; DB 6; Length 469;
Conservative 100.0%; Pred. No. 0.11;
              0; Mismatches   0; Indels   0; Gaps   0;

```
TCATTGTTAGACTTTGAAATTTC 26
||||||||||||||||||||||
TCATTGTTAGACTTTGAAATTTC 425
```

ndard; cDNA; 483 BP.

(first entry)

cancer-associated cDNA, SEQ ID No 558.

cancer; immunogenic; vaccine; tumour; gene; ss.

A2.

2001WO-US023826.

2000US-0223265P.
2000US-0237406P.
2001US-0277495P.
2001US-0302702P.

KA CORP.

J, Secrist H;

7462/30.

cleotide encoding colon tumor polypeptides, useful as
treating colon cancers.

a 309; 425pp; English.

n relates to isolated polynucleotides (I) encoding colon
eptides (II). (I) is useful for stimulating an immune
a patient and treating colon cancer in a patient.
ides derived from (I) are useful for determining the presence
a patient. (I) and (II) are useful in pharmaceutical
, e.g. vaccines, and other compositions for the diagnosis and
colon cancer. A composition comprising a first component
m physiologically acceptable carriers and immunostimulants,
en-presenting cell expressing (II) is useful for inhibiting
of cancer in a patient. (II) is useful in the design and
of ribozyme molecules for inhibiting expression of tumour
and (I). ABK54531-ABK55464 represent human colon cancer cDNA
the invention

BP; 155 A; 97 C; 112 G; 119 T; 0 U; 0 Other;

              1.9%; Score 26; DB 6; Length 483;

Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 26; Conservative 0; Mismatches 0; Indels 0;

```
Qy     1 ATGTCATTGTTAGACTTTGAAATTTC 26
         ||||||||||||||||||||||||||
Db   397 ATGTCATTGTTAGACTTTGAAATTTC 422
```

RESULT 48
AAZ51563
ID   AAZ51563 standard; cDNA; 486 BP.
XX
AC   AAZ51563;
XX
DT   21-JUN-2000 (first entry)
XX
DE   Human hypoxia response regulating gene, 77H4 related cDNA clone
XX
KW   Hypoxia response regulating gene; gene 77H4; human; EST 3D; angi
KW   cardiant; apoptosis; vasotropic; cytostatic; ophthalmological; s
KW   cerebroprotective; antagonist; regulator; inhibitor; treatment;
KW   hypoxia associated pathology; HAP; gene therapy; diagnosis; reti
KW   steroid receptor coactivator; SRA; ischaemia; myocardial infarct
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   polyA_signal 449..454
FT   /*tag= a
XX
PN   WO20001525-A1.
XX
PD   09-MAR-2000.
XX
PF   27-AUG-1999; 99WO-US020394.
XX
PR   27-AUG-1998; 98US-0098158P.
PR   05-MAY-1999; 99US-0132684P.
XX
PA   (QUAR-) QUARK BIOTECH INC.
PA   (KOHN/) KOHN K.
XX
PI   Einat P, Skaliter R, Feinstein E;
XX
DR   WPI; 2000-256577/22.
XX
PT   Novel polynucleotides capable of regulating angiogenesis or apop
PT   useful for diagnosis and treatment of hypoxia, ischemia and tumo
XX
PS   Claim 1; Fig 7b; 78pp; English.
XX
CC   The present sequence is the human hypoxia response regulating ge
CC   related cDNA clone 3D. The gene 77H4 has similarity to steroid r
CC   transcriptional co-activator, SRA function and can serve as a co
CC   in some transcriptional complexes. It has vasotropic, cardiant,
CC   ophthalmological, cytostatic and cerebroprotective activity. Ant
CC   of the encoded protein, functions as a regulator of apoptosis or
CC   angiogenesis. The protein encoded by this polynucleotide, the
CC   biologically active product from enzymatic activity of the prote
CC   inhibitor of the enzymatic activity is useful for regulating hyp
CC   associated pathologies (HAP). It is useful for gene therapy, dia
CC   and treatment of tumour growth and ischaemia, e.g., retinopathy,
CC   myocardial infarction and stroke
XX
SQ   Sequence 486 BP; 160 A; 92 C; 113 G; 121 T; 0 U; 0 Other;

Query Match   1.9%; Score 26; DB 3; Length 486;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 26; Conservative 0; Mismatches 0; Indels 0;

```
Qy     1 ATGTCATTGTTAGACTTTGAAATTTC 26
         ||||||||||||||||||||||||||
Db   400 ATGTCATTGTTAGACTTTGAAATTTC 425
```

lard; DNA; 487 BP.
XX
(first entry)
XX
d to diagnose liver cancer.
XX
ancer; ds; hepatocellular carcinoma; hepatotropic;
ver tumour; cytostatic; expression profile; disease state;
ssion; drug toxicity; drug efficacy; drug metabolism.
XX

2.
XX
2001WO-US030589.
XX
2000US-0237054P.
XX
LOGIC INC.
XX
ares C, Peres-Da-Silva S, Vockley JG;
XX
119/45.
XX
D NO 729; 298pp; English.
XX
d detecting the progression of liver cancer, hepatocellular
metastatic liver tumor in a patient, involves detecting the
ssion of two or more genes in a liver tissue sample.

relates to a novel method for diagnosing and detecting the
liver cancer, hepatocellular carcinoma or metastatic liver
tient, and differentiating metastatic liver cancer from
carcinoma in a patient, involving detecting the level of
two or more genes represented in ABN93503-ABN97455 in a
The method of the invention has hepatotropic, and
ivity. The method is useful for diagnosing and detecting
on of liver cancer, hepatocellular carcinoma and metastatic
na in a patient. The method is useful for identifying
ofiles which serve as useful diagnostic markers as well as
can be used to monitor disease states, disease progression,
drug efficacy and drug metabolism. Note: The sequence data
t did not form part of the printed specification, but was
ectronic format directly from WIPO at
pub/published_pct_sequences
XX
BP; 141 A; 97 C; 118 G; 131 T; 0 U; 0 Other;

rity   1.9%;   Score 26; DB 6; Length 487;
        100.0%;   Pred. No. 0.11;
nservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

ATTGTTAGACTTTGAAATTTC 26
|||||||||||||||||||||
ATTGTTAGACTTTGAAATTTC 452

lard; cDNA; 498 BP.
XX
first entry)

DE   Human endothelial cell cDNA #4820.
XX
KW   Human; ss; sequencing by hybridisation; SBH; expressed sequence t
KW   genome mapping; biodiversity; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   US2003073623-A1.
XX
PD   17-APR-2003.
XX
PF   30-JUL-2001; 2001US-00918995.
XX
PR   30-JUL-2001; 2001US-00918995.
XX
PA   (DRMA/) DRMANAC R T.
PA   (LABA/) LABAT I.
PA   (STAC/) STACHE-CRAIN B.
PA   (DICK/) DICKSON M C.
PA   (JONE/) JONES L W.
XX
PI   Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX
DR   WPI; 2003-615964/58.
XX
PT   New polynucleotide sequences obtained from various cDNA libraries
PT   as hybridization probes, as oligomers for PCR, for chromosome and
PT   mapping, in the recombinant production of protein, or in generati
PT   antisense DNA or RNA.
XX
PS   Claim 1; SEQ ID NO 23899; 44pp; English.
XX
CC   The invention relates to an isolated polynucleotide comprising an
CC   38043 cDNA sequences, appearing as ACH12789-ACH50831, whose seque
CC   determined by the technique of SBH (sequencing by hybridisation).
CC   included is a purified polypeptide comprising a sequence correspo
CC   a reading frame of the novel polynucleotide. The nucleic acid seq
CC   are useful in diagnostics as expressed sequence tags (EST) for
CC   identifying expressed genes or for physical mapping of the human
CC   in forensics, in assessing biodiversity, or in identifying muta
CC   responsible for genetic disorders and other traits. The nucleotid
CC   sequences are also useful as hybridisation probes, as oligomers f
CC   for chromosome and gene mapping, in the recombinant production of
CC   protein, or in generating antisense DNA or RNA. The purified poly
CC   is useful for generating antibodies specific for it. The present
CC   is one of the 38043 isolated cDNA/EST sequences. Note: The sequen
CC   for this patent did not form part of the printed specification, b
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ   Sequence 498 BP; 152 A; 119 C; 117 G; 107 T; 0 U; 3 Other;

Query Match   1.9%; Score 26; DB 8; Length 498;
Best Local Similarity   100.0%; Pred. No. 0.11;
Matches   26; Conservative   0; Mismatches   0; Indels   0; G

Qy   1 ATGTCATTGTTAGACTTTGAAATTTC 26
              ||||||||||||||||||||||||||
Db   349 ATGTCATTGTTAGACTTTGAAATTTC 374

RESULT 51
AAC01272
ID   AAC01272 standard; cDNA; 516 BP.
XX
AC   AAC01272;
XX
DT   06-OCT-2000 (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 1270.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isol
KW   gene therapy; chromosome mapping; ss.

```
PR    03-JUL-2001; 2001US-0302702P.
XX
PA    (CORI-) CORIXA CORP.
XX
PI    Pyle RA, Xu J, Secrist H;
XX
DR    WPI; 2002-257462/30.
XX
PT    Novel polynucleotide encoding colon tumor polypeptides, useful as
PT    vaccines for treating colon cancers.
XX
PS    Claim 1; Page 403; 425pp; English.
XX
CC    The invention relates to isolated polynucleotides (I) encoding co
CC    tumour polypeptides (II). (I) is useful for stimulating an immune
CC    response in a patient and treating colon cancer in a patient.
CC    Oligonucleotides derived from (I) are useful for determining the
CC    of cancer in a patient. (I) and (II) are useful in pharmaceutical
CC    compositions, e.g. vaccines, and other compositions for the diagr
CC    treatment of colon cancer. A composition comprising a first compo
CC    selected from physiologically acceptable carriers and immunostim
CC    and an antigen-presenting cell expressing (II) is useful for inhi
CC    development of cancer in a patient. (I) is useful in the design a
CC    preparation of ribozyme molecules for inhibiting expression of tu
CC    polypeptides and (I). ABK54531-ABK55464 represent human colon car
CC    sequences of the invention
XX
SQ    Sequence 531 BP; 153 A; 105 C; 128 G; 143 T; 0 U; 2 Other;

Query Match     1.9%;    Score 26;  DB 6;   Length 531;
Best Local Similarity   100.0%;  Pred. No. 0.11;
Matches  26;  Conservative   0;  Mismatches   0;  Indels   0;  C

Qy     1 ATGTCATTGTTAGACTTTGAAATTTC 26
         ||||||||||||||||||||||||||
Db   459 ATGTCATTGTTAGACTTTGAAATTTC 484


RESULT 53
AAC01271
ID   AAC01271 standard; cDNA; 540 BP.
XX
AC   AAC01271;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 1269.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isc
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-00200610.
XX
PR   26-FEB-1999;  99US-0122487P.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR   WPI; 2000-500381/45.
DR   P-PSDB; AAG01265.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and fo
PT   diagnostic, forensic, gene therapy and chromosome mapping procedu
XX
PS   Claim 1; SEQ ID NO 1269; 71pp + Sequence Listing; English.
```

```
2000EP-00200610.
99US-0122487P.
T.
dwards J, Duclert A, Giordano J;
381/45.
266.
cid that is a 5' expressed sequence tag (5' EST) for
As and genomic DNAs that correspond to 5'ESTs and for
orensic, gene therapy and chromosome mapping procedures.
ID NO 1270; 71pp + Sequence Listing; English.
equence is one of a large number of 5' ESTs derived from
g secreted proteins. An ORF has been identified within the
5' ESTs were prepared from total human RNAs or polyA+ RNAs
30 different tissues. EST sequences usually correspond
3' untranslated region (UTR) of the mRNA because they are
d from oligo-dT primed cDNA libraries. Such ESTs are not
n in those cases where longer cDNA sequences have been
full 5' UTR is rarely included. 5' ESTs are derived from
tact 5' ends and can therefore be used to obtain full length
omic DNAs. 5' ESTs are also used in diagnostic, forensic,
and chromosome mapping procedures. They are used to obtain
latory sequences and to design expression and secretion
BP; 134 A; 106 C; 131 G; 145 T; 0 U; 0 Other;
arity   1.9%;   Score 26;  DB 3;   Length 516;
        100.0%;  Pred. No. 0.11;
onservative   0;  Mismatches   0;  Indels   0;  Gaps   0;
CATTGTTAGACTTTGAAATTTC 26
||||||||||||||||||||||
CATTGTTAGACTTTGAAATTTC 503
dard; cDNA; 531 BP.
(first entry)
ancer-associated cDNA, SEQ ID No 877.
cancer; immunogenic; vaccine; tumour; gene; ss.
2.
2001WO-US023826.
2000US-0223265P.
2000US-0237406P.
2001US-0277495P.
```

equence is one of a large number of 5' ESTs derived from
g secreted proteins. An ORF has been identified within the
5' ESTs were prepared from total human RNAs or polyA+ RNAs
30 different tissues. EST sequences usually correspond
3' untranslated region (UTR) of the mRNA because they are
d from oligo-dT primed cDNA libraries. Such ESTs are not
or isolating cDNA sequences derived from the 5' ends of
n in those cases where longer cDNA sequences have been
full 5' UTR is rarely included. 5' ESTs are derived from
tact 5' ends and can therefore be used to obtain full length
omic DNAs. 5' ESTs are also used in diagnostic, forensic,
and chromosome mapping procedures. They are used to obtain
latory sequences and to design expression and secretion

BP; 135 A; 110 C; 137 G; 152 T; 0 U; 6 Other;

arity   1.9%;   Score 26;   DB 3;   Length 540;
onservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

CATTGTTAGACTTTGAAATTTC 26
|||||||||||||||||||||||
CATTGTTAGACTTTGAAATTTC 527

lard; cDNA; 570 BP.

(first entry)

ssociated gene sequence SEQ ID NO:292.

associated gene; cancer antigen; detection; cancer;
ostatic; proliferative; vulnerary; immunomodulator;
antiasthmatic; antirheumatic; antiarthritic; antiviral;
ry; antithyroid; antiallergic; antibacterial; cardiant;
; neuroprotective; thrombolytic; coagulant; nootropic;
tipsoriatic; antiangiogenic; gene therapy; inflammation;
r; haematopoietic cell disorder; autoimmune disorder;
ion; graft versus host disease; organ rejection;
hrombolytic; cardiovascular disorder; infection;
isease; drug screening; ss.

000WO-US005882.

99US-0124270P.

GENOME SCI INC.

en SM;

33/55.
89.

nucleic acids comprising sequences encoding peptides
ating or diagnosing e.g. cancer.

853; 2352pp; English.

C78448 encode the human cancer associated proteins given in
B44239. The proteins can have activities based on the
118 the genes are expressed in. Example of activities

CC   include: cytostatic; proliferative; vulnerary; immunomodulator;
CC   antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC   antiinflammatory; antithyroid; antiallergic; antibacterial; antiv
CC   dermatological; neuroprotective; cardiant; thrombolytic; coagular
CC   nootropic; vasotropic; antipsoriatic and antiangiogenic. The
CC   polynucleotides and polypeptides can be used for preventing, trea
CC   ameliorating medical conditions and diagnosing pathological condi
CC   Polynucleotides, polypeptides, antibodies, agonists and antagonis
CC   the present invention may be used to treat immune disorders by ac
CC   or inhibiting the proliferation, differentiation or mobilisation
CC   immune cells, to treat disorders of haematopoietic cells, autoimm
CC   disorders, allergic reactions, graft versus host disease and orga
CC   rejection, modulate haemostatic or thrombolytic activity, modulat
CC   inflammation, cancers, cardiovascular disorders, neurological dis
CC   bacterial or viral infections. The peptides, nucleotides, antibod
CC   agonists and antagonists may be also be used in drug screens. AAC
CC   AAC78457 and AAB44240 represent sequences used in the exemplifica
CC   the present invention
XX
SQ   Sequence 570 BP; 171 A; 112 C; 139 G; 145 T; 0 U; 3 Other;

Query Match   1.9%;   Score 26;   DB 3;   Length 570;
Best Local Similarity   100.0%;   Pred. No. 0.11;
Matches 26;   Conservative   0;   Mismatches   0;   Indels   0;   G

Qy   1 ATGTCATTGTTAGACTTTGAAATTTC 26
         ||||||||||||||||||||||||||
Db   469 ATGTCATTGTTAGACTTTGAAATTTC 494

RESULT 55
AA251562
ID   AA251562 standard; cDNA; 580 BP.
XX
AC   AA251562;
XX
DT   21-JUN-2000   (first entry)
XX
DE   Human hypoxia response regulating gene, 77H4 related cDNA clone 1
XX
KW   Hypoxia response regulating gene; gene 77H4; human; EST 18E; card
KW   apoptosis; angiogenesis; vasotropic; cytostatic; ophthalmological
KW   cerebroprotective; antagonist; regulator; inhibitor; treatment; t
KW   hypoxia associated pathology; HAP; gene therapy; diagnosis; ischa
KW   steroid receptor coactivator; SRA; retinopathy; myocardial infarc
KW   stroke; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   polyA_signal   536..541
FT   /*tag= a
XX
PN   WO200012525-A1.
XX
PD   09-MAR-2000.
XX
PF   27-AUG-1999;   99WO-US020394.
XX
PR   27-AUG-1998;   98US-0098158P.
PR   05-MAY-1999;   99US-013268 4P.
XX
PA   (QUAR-) QUARK BIOTECH INC.
PA   (KOHN/) KOHN K.
XX
PI   Einat P, Skaliter R, Feinstein E;
XX
DR   WPI; 2000-256577/22.
XX
PT   Novel polynucleotides capable of regulating angiogenesis or apopto
PT   useful for diagnosis and treatment of hypoxia, ischemia and tumor
XX
PS   Claim 1; Fig 7a; 78pp; English.

CC subjected to neurotoxic stress. The invention can be used to pro
CC neural cells from, and ameliorate the effects of, hypoxia and is
CC and thus in the treatment of stroke, hypoxia and ischaemia. The
CC is also useful to prevent apoptosis and promote angiogenesis. Th
CC invention can be used in diagnostic assays for cells that have b
CC subjected to hypoxia or ischaemia, and in screening assays to id
CC agents capable of enhancing gene expression. The present nucleic
CC sequence represents the human gene 7744 gene of the invention. N
CC specification states that this sequence encodes the human gene 7
CC protein (ABG71808) but this does not appear to be the case
XX
SQ Sequence 580 BP; 177 A; 111 C; 134 G; 158 T; 0 U; 0 Other;

Query Match   1.9%; Score 26; DB 7; Length 580;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 26; Conservative 0; Mismatches 0; Indels 0;

Qy   1 ATGTCATTGTTAGACTTTGAAATTTC 26
          ||||||||||||||||||||||||||
Db   487 ATGTCATTGTTAGACTTTGAAATTTC 512

RESULT 57
ABQ58318/c
ID ABQ58318 standard; cDNA; 400 BP.
XX
AC ABQ58318;
XX
DT 02-AUG-2002 (first entry)
XX
DE Human colon cancer related nucleotide sequence SEQ ID NO:2013.
XX
KW Human; colon cancer; cancer; tissue profiling; forensic; mapping
KW genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200229086-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US030732.
XX
PR 02-OCT-2000; 2000US-0237271P.
XX
PA (FARB) BAYER CORP.
XX
PI Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molin(
PI Thiaglingam A, Lewis ME;
XX
DR WPI; 2002-426115/45.
XX
PT New isolated nucleic acid that is differentially expressed in ca
PT tissues useful for determining the presence of colon cancer in a
PT tissue type, and in antisense therapy.
XX
PS Claim 1; Fig 1; 796pp; English.
XX
CC ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differ
CC expressed in cancer tissues. ABB78993 to ABB79004 represent prot
CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I)
CC used in antisense therapy. An antibody immunoreactive with a pol
CC encoded by (I) is useful for detecting cancer in a patient sampl
CC for detecting the presence or absence of a polynucleotide encode
CC nucleic acid which hybridises to (I) in a cell. A probe/primer d
CC from (I) can be used for determining the presence of a nucleic a
CC hybridises to (I), and for determining the phenotype of cells in
CC of cells from a patient. (I) is useful for determining the prese
CC colon cancer in a cell or tissue type, for determining the prese
CC state of other type of cancer, in antisense therapy, to generate
CC macroarrays on a solid surface, to identify a chromosome on which
CC corresponding gene resides, and in tissue profiling, forensics,
CC analysis, mapping and diagnostic applications. (I) can be used tc

---

sequence is the human hypoxia response regulating gene, 77H4,
clone 18E. The gene 77H4 has similarity to steroid receptor
nal co-activator, SRA function and can serve as a coactivator
scriptional complexes. It has vasotropic, cardiant,
ical, cytostatic and cerebroprotective activity. Antagonist
ed protein, functions as a regulator of apoptosis or
The protein encoded by this polynucleotide, the
active product from enzymatic activity of the protein or
the enzymatic activity is useful for regulating hypoxia
athologies (HAP). It is useful for gene therapy, diagnosis
of tumour growth and ischaemia, e.g., retinopathy,
nfarction and stroke

BP; 177 A; 111 C; 134 G; 158 T; 0 U; 0 Other;

larity 100.0%; Pred. No. 0.11;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[CATTGTTAGACTTTGAAATTTC 26
||||||||||||||||||||||
CATTGTTAGACTTTGAAATTTC 512

dard; cDNA; 580 BP.

(first entry)

744 gene sequence.

ta-regulated activity; neurotoxic stress; hypoxia; ischaemia;
osis; angiogenesis; cerebroprotective; gene therapy;
inhibitor of oxidative stress-mediated apoptosis;
angiogenesis; gene 7744; gene; ss.

.A1.

2001US-00802472.

97US-0056453P.
98US-0013810S.
98US-0098158P.
99US-0132684P.
99US-00384096.

.P.
TER R.
STEIN E.

liter R, Feinstein E;

8474/02.

d polypeptides and genes associated with hypoxia-regulated
ful for treating stroke, hypoxia and ischemia.

: 61; 72pp; English.

nvention relates to a new polypeptide associated with
ated activity. The invention is useful in diagnostic assays.
is further useful as a diagnostic tool which can be used to
ir presence in a cell. The invention is also useful for
odies that could be used in diagnostic assays for the
he protein and for determining if any given cell had been

nd to screen for peptide analogues and antagonists

BP; 100 A; 80 C; 61 G; 138 T; 0 U; 21 Other;

arity 1.8%; Score 25; DB 6; Pred. No. 0.32; Length 400;
onservative 0; Mismatches 0; Indels 0; Gaps 0;

```
      CATTGTTAGACTTTGAAATTT 25
      |||||||||||||||||||||
      CATTGTTAGACTTTGAAATTT 69
```

dard; cDNA; 626 BP.

(first entry)

novel human diagnostic protein #230.

some mapping; gene mapping; gene therapy; forensic;
nt; medical imaging; diagnostic; genetic disorder; ss.

2.

2001WO-US008631.

2000US-00540217.
2000US-00649167.

INC.

Liu C, Tang YT;

362/73.
239.

polynucleotide and encoded polypeptides, useful in
orensics, gene mapping, identification of mutations
or genetic disorders or other traits and to assess

ID NO 230; 103pp; English.

relates to isolated polynucleotide (I) and polypeptide (II)
is useful as hybridisation probes, polymerase chain
primers, oligomers, and for chromosome and gene mapping,
nant production of (II). The polynucleotides are also used
as expressed sequence tags for identifying expressed
useful in gene therapy techniques to restore normal
I) or to treat disease states involving (II). (II) is
nerating antibodies against it, detecting or quantitating a
tissue, as molecular weight markers and as a food
I) and its binding partners are useful in medical imaging
ssing (II). (I) and (II) are useful for treating disorders
rant protein expression or biological activity. The
nd polynucleotide sequences have applications in
orensics, gene mapping, identification of mutations
or genetic disorders or other traits to assess biodiversity
other types of data and products dependent on DNA and
uences. AAS64197-AAS94564 represent novel human diagnostic
es of the invention. Note: The sequence data for this
appear in the printed specification, but was obtained in
mat directly from WIPO at
pub/published_pct_sequences

SQ　Sequence 626 BP; 173 A; 126 C; 157 G; 170 T; 0 U; 0 Other;

Query Match　1.8%; Score 25; DB 5; Length 626;
Best Local Similarity　100.0%; Pred. No. 0.31;
Matches　25; Conservative　0; Mismatches　0; Indels　0; C

```
Qy      1 ATGTCATTGTTAGACTTTGAAATTT 25
          |||||||||||||||||||||||||
Db    542 ATGTCATTGTTAGACTTTGAAATTT 566
```

RESULT 59
AAX23451/c
ID　AAX23451 standard; DNA; 24 BP.
XX
AC　AAX23451;
XX
DT　18-JUN-1999　(first entry)
XX
DE　Human TNRL3 RACE primer 2.
XX
KW　Tumour necrosis factor receptor; signal transducer molecule; TNF;
KW　developmental abnormality; gestational abnormality; prostate ca
KW　APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy;
KW　cytoplasmic domain; immunogen; antibody preparation; breast carci
KW　apoptosis; human; primer; ss.
XX
OS　Synthetic.
OS　Homo sapiens.
XX
PN　WO9911791-A2.
XX
PD　11-MAR-1999.
XX
PF　04-SEP-1998;　98WO-US018393.
XX
PR　05-SEP-1997;　97US-00924634.
XX
PA　(UNIW) UNIV WASHINGTON.
XX
PI　Chaudhary PM;
XX
DR　WPI; 1999-205197/17.
XX
PT　New Tumor Necrosis Factor family receptor polypeptides and ligand
PT　useful for diagnosis and treatment of prostate cancer and develop
PT　or gestational abnormalities.
XX
PS　Example VII; Page 121; 156pp; English.
XX
CC　This invention describes isolated Tumor Necrosis Factor (TNF) fam
CC　receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC　fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TN
CC　their active fragments. APO4 is useful for diagnosing prostate ca
CC　determining levels of APO4 in an individual. Prostate cancer can
CC　treated using APO4 selective binding agents linked to a therapeut
CC　moiety. APO4 polypeptides are also useful for identifying selecti
CC　binding agents, useful in diagnosis/treatment of disease by bindi
CC　agents to the polypeptide/active fragment which is extracellular,
CC　expressed on the cell surface. The binding is preferably performe
CC　vivo. APO4 polypeptides/ active fragments are also useful for scr
CC　for agonists and antagonists by binding and observing the changer
CC　activity. Effective pharmacological agents useful in diagnosis or
CC　treatment of disease are also identified using APO4 polypeptides/
CC　fragments and APO4 signal transducer molecules that specifically
CC　with a cytoplasmic domain of APO4 and detecting a change in level
CC　activity. The method is performed in vivo or in vitro. APO polype
CC　are all useful as immunogens for preparing antibodies. APO4 is al
CC　useful for diagnosis/treatment of developmental or gestational
CC　abnormalities. APO8 was transfected to human breast carcinoma cel
CC　MCF-7, and induced apoptosis
XX
SQ　Sequence 24 BP; 7 A; 9 C; 3 G; 5 T; 0 U; 0 Other;

larity 1.7%; Score 24; DB 2; Length 24;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
GATGAGGGGAAGGCTGTCTAC 633
||||||||||||||||||||
GATGAGGGGAAGGCTGTCTAC 1
```

ndard; DNA; 24 BP.

(first entry)

RACE primer 1.

sis factor receptor; signal transducer molecule; TNF; APO4;
l abnormality; gestational abnormality; prostate cancer;
APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
domain; immunogen; antibody preparation; breast carcinoma;
uman; primer; ss.

98WO-US018393.

97US-00924634.

WASHINGTON.

5191/17.

crosis Factor family receptor polypeptides and ligands -
iagnosis and treatment of prostate cancer and developmental
al abnormalities.

Page 121; 156pp; English.

on describes isolated Tumor Necrosis Factor (TNF) family
ypeptides: APO4, APO6, APO8 and APO9 or their active
nd isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
fragments. APO4 is useful for diagnosing prostate cancer by
levels of APO4 in an individual. Prostate cancer can also be
g APO4 selective binding agents linked to a therapeutic
polypeptides are also useful for identifying selective
s, useful in diagnosis/treatment of disease by binding of
a polypeptide/active fragment which is extracellular, or
the cell surface. The binding is preferably performed in
olypeptides/ active fragments are also useful for screening
and antagonists by binding and observing the changer in APO4
ective pharmacological agents useful in diagnosis or
disease are also identified using APO4 polypeptides/active
d APO4 signal transducer molecules that specifically interact
lasmic domain of APO4 and detecting a change in level of APO4
a method is performed in vivo or in vitro. APO polypeptides
l as immunogens for preparing antibodies. APO4 is also
iagnosis/treatment of developmental or gestational
5. APO8 was transfected to human breast carcinoma cell line
duced apoptosis

BP; 7 A; 2 C; 13 G; 2 T; 0 U; 0 Other;

Query Match 1.7%; Score 24; DB 2; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 24; Conservative 0; Mismatches 0; Indels 0;

```
Qy    812 CTGCCCCTTCCTCACCTACTTCG 835
          ||||||||||||||||||||||||
Db     24 CTGCCCCTTCCTCACCTACTTCG 1
```

RESULT 61
AAX56003
ID   AAX56003 standard; DNA; 40 BP.
XX
AC   AAX56003;
XX
XX
DT   15-JUL-1999 (first entry)
XX
DE   Human tumour necrosis factor Apo-3 ligand PCR primer SEQ ID NO:5
XX
KW   Human; tumour necrosis factor; Apo-3 ligand; lymphotoxin; apopto
KW   NF-kappaB-dependent transcription; JNK/SAPK-dependent response;
KW   PCR primer; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9919490-A1.
XX
PD   22-APR-1999.
XX
PF   09-OCT-1998;   98WO-US021407.
XX
PR   10-OCT-1997;   97US-0062037P.
PR   17-DEC-1997;   97US-0069862P.
XX
PA   (GETH) GENENTECH INC.
XX
PI   Ashkenazi AJ, Marsters SA, Pitti R;
XX
DR   WPI; 1999-287982/24.
XX
PT   New human Apo3- ligand (a tumor necrosis factor) homologue.
XX
PS   Example 2; Page 37; 74pp; English.
XX
CC   The present invention describes a human tumour necrosis factor (
CC   lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has
CC   cytostatic activity. Apo-3 ligand can be used to induce apoptosi
CC   mammalian cancer cells, to induce NF-kappaB-dependent transcript
CC   to induce JNK/SAPK-dependent responses in mammalian cells. The p
CC   sequence represents an Apo-3 ligand PCR primer, which is used in
CC   example from the present invention
XX
SQ   Sequence 40 BP; 9 A; 13 C; 13 G; 5 T; 0 U; 0 Other;

Query Match 1.7%; Score 24; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 24; Conservative 0; Mismatches 0; Indels 0;

```
Qy    244 CGGGCATCGCTGTCCGCCCAGGAG 267
          ||||||||||||||||||||||||
Db     17 CGGGCATCGCTGTCCGCCCAGGAG 40
```

RESULT 62
ABK40355
ID   ABK40355 standard; DNA; 23 BP.
XX
AC   ABK40355;
XX
DT   15-JUL-2002 (first entry)
XX
DE   Probe for gene amplification analysis of human PRO207.

enign tumour; malignant tumour; lymphoid malignancy;
uronal disorder; stromal disorder; blastocoelic disorder;
disorder; immune disorder; angiogenic disorder; cytostatic;
ve; probe; ss.

1.

2000WO-US003565.

99WO-US005028.
99US-0123972P.
99US-0133459P.
99WO-US012252.
99US-0140650P.
99US-0140653P.
99US-0144758P.
99US-0145698P.
99US-0146222P.
99US-0149395P.
99US-0151689P.
99WO-US020111.
99WO-US021090.
99WO-US028313.
99WO-US028301.
2000WO-US000219.

TECH INC.

Goddard A, Godowski PJ, Gurney AL, Hillan KJ,
Pan J, Pitti RM, Roy MA, Smith V, Stone DM;
Wood WI;

567/26.

cleic acids encoding PRO polypeptides, useful for treating
ignant tumors, leukemias and lymphoid malignancies,
angiogenic and immunologic disorders.

ge 140; 302pp; English.

vention relates to the isolation of novel human PRO
(AAU86128-AAU86162) and the polynucleotide sequences
The PRO polypeptides, agonists, antagonists or anti-PRO
useful for treating benign or malignant tumours (e.g.
bladder, breast, etc), leukaemias and lymphoid
other disorders such as neuronal, glial, astrocytal,
glandular, macrophagal, stromal and blastocoelic disorders,
immune and angiogenic disorders. The polynucleotide
also useful in gene therapy. The present sequence
robe used in the methods of the present invention

; 1 A; 7 C; 7 G; 8 T; 0 U; 0 Other;

rity   1.7%;   Score 23;   DB 6;   Length 23;
100.0%;   Pred. No. 3.1;
nservative   0; Mismatches   0; Indels   0; Gaps   0;

TGGGCCTGTTCACGTGTT 1003
||||||||||||||||||
TGGGCCTGTTCACGTGTT 23

ard; DNA; 38 BP.

XX
DT   15-JUL-1999 (first entry)
XX
DE   Human tumour necrosis factor Apo-3 ligand PCR primer SEQ ID NO:6.
XX
KW   Human; tumour necrosis factor; Apo-3 ligand; lymphotoxin; apopto;
KW   NF-kappaB-dependent transcription; JNK/SAPK-dependent response; c
KW   PCR primer; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9919490-A1.
XX
PD   22-APR-1999.
XX
PF   09-OCT-1998;   98WO-US021407.
XX
PR   10-OCT-1997;   97US-0062037P.
PR   17-DEC-1997;   97US-0069862P.
XX
PA   (GETH) GENENTECH INC.
XX
PI   Ashkenazi AJ, Marsters SA, Pitti R;
XX
DR   WPI; 1999-287982/24.
XX
PT   New human Apo3- ligand (a tumor necrosis factor) homologue.
XX
PS   Example 2; Page 37; 74pp; English.
XX
CC   The present invention describes a human tumour necrosis factor (T
CC   lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has
CC   cytostatic activity. Apo-3 ligand can be used to induce apoptosis
CC   mammalian cancer cells, to induce NF-kappaB-dependent transcripti
CC   to induce JNK/SAPK-dependent responses in mammalian cells. The pr
CC   sequence represents an Apo-3 ligand PCR primer, which is used in
CC   example from the present invention
XX
SQ   Sequence 38 BP; 9 A; 11 C; 12 G; 6 T; 0 U; 0 Other;

Query Match   1.6%; Score 22; DB 2;   Length 38;
Best Local Similarity   100.0%; Pred. No. 8.9;
Matches 22; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   834 CGGACTCTTCCAGGTTCACTGA 855
            ||||||||||||||||||||||
Db   38 CGGACTCTTCCAGGTTCACTGA 17

RESULT 64
AAL18265
ID   AAL18265 standard; cDNA; 223 BP.
XX
AC   AAL18265;
XX
DT   07-DEC-2001 (first entry)
XX
DE   Human breast cancer expressed polynucleotide 10722.
XX
KW   Human; breast cancer; cell marker; cytostatic; ss.
XX
OS   Homo sapiens.
XX
PN   WO200151628-A2.
XX
PD   19-JUL-2001.
XX
PF   10-JAN-2001; 2001WO-US000798.
XX
PR   14-JAN-2000; 2000US-0176077P.
PR   14-MAR-2000; 2000US-0189167P.
PR   24-MAR-2000; 2000US-0192099P.

2000US-0193480P.
2000US-0205230P.
2000US-0211315P.
2000US-0220534P.

ENNIUM PREDICTIVE MEDICINE INC.

u Y, Wang Y, Steinmann K;

1856/48.

useful as a marker for the diagnosis of breast cancer.

e 1912-1913; 3695pp; English.

n relates to human breast cancer expressed polynucleotides
L26789) and methods of assessing whether a patient is
th breast cancer by examining the correlation between the
f certain markers and the cancerous state of breast cells.
eotides and encoded polypeptides are potential markers for
diagnosing, monitoring, characterising treating and
preventing breast cancer. The polynucleotides and encoded
are also useful for isolating compounds with cytostatic

BP; 46 A; 51 C; 48 G; 78 T; 0 U; 0 Other;

larity   1.6%;   Score 22;   DB 4;   Length 223;
Conservative   100.0%;   Pred. No. 8.1;
               0; Mismatches   0; Indels   0; Gaps   0;

TATTATTTTTATTATTATT 1351
||||||||||||||||||
TATTATTTTTATTATTATT 119

ndard; cDNA; 263 BP.

A2.

(first entry)

cancer expressed polynucleotide 1278.

t cancer; cell marker; cytostatic; ss.

2001WO-US000798.

2000US-0176077P.
2000US-0189167P.
2000US-0192099P.
2000US-0193480P.
2000US-0205230P.
2000US-0211315P.
2000US-0220534P.

ENNIUM PREDICTIVE MEDICINE INC.

J Y, Wang Y, Steinmann K;

1856/48.

useful as a marker for the diagnosis of breast cancer.

a 278; 3695pp; English.

XX
CC The invention relates to human breast cancer expressed polynucle
CC (AAL07544-AAL26789) and methods of assessing whether a patient
CC afflicted with breast cancer by examining the correlation betwe
CC expression of certain markers and the cancerous state of breast
CC The polynucleotides and encoded polypeptides are potential marke
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and er
CC polypeptides are also useful for isolating compounds with cytost
CC activity
XX
SQ Sequence 263 BP; 52 A; 63 C; 63 G; 85 T; 0 U; 0 Other;

Query Match   1.6%;   Score 22;   DB 4;   Length 263;
Best Local Similarity 100.0%;   Pred. No. 8.1;
Matches 22; Conservative   0; Mismatches   0; Indels   0;

Qy   1330 AGATATTATTTTTATTATTATT 1351
            ||||||||||||||||||||||
Db   137 AGATATTATTTTTATTATTATT 158

RESULT 66
ADA02995
ID   ADA02995 standard; cDNA; 1005 BP.
XX
AC   ADA02995;
XX
DT   06-NOV-2003 (first entry)
XX
DE   Mouse Sept9 carcinoma associated coding sequence, SEQ ID NO:1513
XX
KW   Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer
KW   prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug sc
KW   gene; ss.
XX
OS   Mus sp.
XX
PN   WO200305146-A2.
XX
PD   17-JUL-2003.
XX
PF   26-DEC-2002; 2002WO-US041414.
XX
PR   26-DEC-2001; 2001US-00035832.
XX
PA   (SAGR-) SAGRES DISCOVERY.
XX
PI   Morris DW;
XX
DR   WPI; 2003-587068/55.
XX
PT   New recombinant nucleic acid encoding carcinoma associated prote
PT   useful for preparing compositions for treating carcinomas.
XX
PS   Claim 1; SEQ ID NO 1513; 245pp; English.
XX
CC   The invention relates to recombinant carcinoma associated (CA) n
CC   acid sequences from mouse and human (ADA01482-ADA03094), and to
CC   recombinant carcinoma associated proteins (CAP) encoded by them.
CC   invention also encompasses expression vectors and host cells com
CC   CA nucleic acid, a polypeptide (especially an antibody) that spe
CC   binds to the protein, and a biochip comprising CA nucleic acid o
CC   fragments thereof. The sequences of the invention were identifie
CC   oncogenic retroviruses, which insert into the genome of the host
CC   at random. Many of these do not carry transduced host oncogenes
CC   pathogenic trans-acting viral genes, meaning that cancer inciden
CC   direct consequence of the effects of proviral integration into h
CC   protooncogenes. The CA nucleic acid sequences can be used to dia
CC   carcinoma (especially breast cancer, prostate cancer, lymphoma o
CC   leukaemia) or a propensity to carcinoma by determination of the
CC   of a CA gene, or by determination of CA gene expression in parti
CC   tissues. CA nucleic acids, proteins and antibodies are also usefu

...gents and in screening and evaluating drug candidates. The
...ence represents a specifically claimed murine CA nucleic acid
...he invention. Note: The complete sequence data for this
...t form part of the printed specification, but was obtained
... format directly from WIPO at
...pub/published_pct_sequences.

... BP; 285 A; 268 C; 273 G; 179 T; 0 U; 0 Other;

...arity 100.0%; Pred. No. 7.6;   Length 1005;
...onservative 0; Mismatches 0; Indels 0; Gaps 0;

...CCCATTATGAAGTTCATC 448
...|||||||||||||||||
...CCCATTATGAAGTTCATC 396

dard; cDNA; 1005 BP.

(first entry)

DNA.

tostatic; gene therapy; vaccine; carcinoma; lymphomas;
asm; adenocarcinoma; sarcoma; gene.

A2.

2001WO-US051291.

2001US-00798586.
2001US-00004113.
2001US-00052482.
2001US-00997722.
2001US-00034650.

S DISCOVERY.

ngelhard EK;

337/23.

t nucleic acid, useful for treating carcinomas, lymphomas,
Lasm, adenocarcinoma, or sarcomas.

ID NO 561; 2304pp; English.

relates to a novel recombinant nucleic acid comprising a
quence selected from any of the 660 sequences fully defined
ication. A polynucleotide of the invention has cytostatic
may have a use in gene therapy, or in a vaccine. The
cleic acids and polypeptides are useful for treating
g. lymphomas, cancers, neoplasm, adenocarcinoma, and
present sequence represents a mouse cDNA of the invention.

BP; 285 A; 268 C; 273 G; 179 T; 0 U; 0 Other;

arity 100.0%; Pred. No. 7.6;   Length 1005;
onservative 0; Mismatches 0; Indels 0; Gaps 0;

CCATTATGAAGTTCATC 448
|||||||||||||||
CCATTATGAAGTTCATC 396

RESULT 68
ADC85475
ID   ADC85475 standard; DNA; 1005 BP.
XX
AC   ADC85475;
XX
DT   01-JAN-2004 (first entry)
XX
DE   Mouse Sept19 coding sequence.
XX
KW   Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated ;
KW   secreted; transmembrane; intracellular; ds.
XX
OS   Mus sp.
XX
PN   WO200305230-A2.
XX
PD   05-JUN-2003.
XX
PF   02-DEC-2002; 2002WO-US038582.
XX
PR   30-NOV-2001; 2001US-00997722.
XX
PA   (SAGR-) SAGRES DISCOVERY.
XX
PI   Morris DW, Engelhard EK;
XX
DR   WPI; 2003-513603/48.
XX
PT   New recombinant nucleic acid comprising a nucleotide sequence of
PT   the carcinoma-associated (CA) genes, useful for screening for dru
PT   candidates for diagnosing or treating carcinomas.
XX
PS   Claim 1; SEQ ID NO 261; 983pp; English.
XX
CC   The invention relates to a recombinant nucleic acid comprising a
CC   nucleotide sequence selected from any of the fully defined carci
CC   associated (CA) genes from the 50 tables given in the specificati
CC   CA proteins are secreted, transmembrane or intracellular proteins
CC   recombinant nucleic acids are useful for screening for drug candi
CC   for diagnosing or treating carcinomas. Sequences given in ADC8521
CC   ADC85514 represent CA genes of the invention.
XX
SQ   Sequence 1005 BP; 285 A; 268 C; 273 G; 179 T; 0 U; 0 Other;

Query Match   1.6%; Score 22; DB 9; Length 1005;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; G

Qy   427 GCAGCCCATTATGAAGTTCATC 448
         ||||||||||||||||||||||
Db   375 GCAGCCCATTATGAAGTTCATC 396

RESULT 69
AAS84907
ID   AAS84907 standard; cDNA; 1778 BP.
XX
AC   AAS84907;
XX
DT   13-FEB-2002 (first entry)
XX
DE   DNA encoding novel human diagnostic protein #20711.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; s:
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX

2001WO-US008631.

2000US-00540217.
2000US-00649167.

Q INC.

Liu C, Tang YT;

9362/73.
0720.

polynucleotide and encoded polypeptides, useful in
forensics, gene mapping, identification of mutations
for genetic disorders or other traits and to assess

ID NO 2071i; 103pp; English.

n relates to isolated polynucleotide (I) and polypeptide (II)
I) is useful as hybridisation probes, polymerase chain
R) primers, oligomers, and for chromosome and gene mapping,
binant production of (II). The polynucleotides are also used
cs as expressed sequence tags for identifying expressed
s useful in gene therapy techniques to restore normal
(II) or to treat disease states involving (II). (II) is
enerating antibodies against it, detecting or quantitating a
in tissue, as molecular weight markers and as a food
(II) and its binding partners are useful in medical imaging
ressing (II). (I) and (II) are useful for treating disorders
errant protein expression or biological activity. The
and polynucleotide sequences have application in
forensics, gene mapping, identification of mutations
for genetic disorders or other traits to assess biodiversity
ce other types of data and products dependent on DNA and
sequences. AAS64197-AAS94564 represent novel human diagnostic
nces of the invention. Note: The sequence data for this
ot appear in the printed specification, but was obtained in
ormat directly from WIPO at
/pub/published_pct_sequences

B BP; 472 A; 408 C; 371 G; 527 T; 0 U; 0 Other;

1.6%;   Score 22;   DB 5;   Length 1778;
larity   100.0%;   Pred. No. 7.3;
Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

GATGGGGGGCGGGTGAGG 84
||||||||||||||||||
GATGGGGGGCGGGTGAGG 205

idard; cDNA; 2942 BP.

(first entry)

carcinoma associated cDNA, SEQ ID NO:1512.

-A2.

a; carcinoma associated; oncogene; carcinoma; cancer; breast;
nphoma; leukaemia; cytostatic; gene therapy; drug screening;

XX
PF   26-DEC-2002; 2002WO-US041414.
XX
PR   26-DEC-2001; 2001US-00035832.
XX
PA   (SAGR-) SAGRES DISCOVERY.
XX
PI   Morris DW;
XX
DR   WPI; 2003-587068/55.
XX
PT   New recombinant nucleic acid encoding carcinoma associated prote
PT   useful for preparing compositions for treating carcinomas.
XX
PS   Claim 1; SEQ ID NO 1512; 245pp; English.
XX
CC   The invention relates to recombinant carcinoma associated (CA) n
CC   acid sequences from mouse and human (ADA01482-ADA03094), and to
CC   recombinant carcinoma associated proteins (CAP) encoded by them.
CC   invention also encompasses expression vectors and host cells com
CC   CA nucleic acid, a polypeptide (especially an antibody) that spe
CC   binds to the protein, and a biochip comprising CA nucleic acid o
CC   fragments thereof. The sequences of the invention were identifie
CC   oncogenic retroviruses, which insert into the genome of the host
CC   at random. Many of these do not carry transduced host oncogenes
CC   pathogenic trans-acting viral genes, meaning that cancer inciden
CC   direct consequence of the effects of proviral integration into h
CC   protooncogenes. The CA nucleic acid sequences can be used to dia
CC   carcinoma (especially breast cancer, prostate cancer, lymphoma o
CC   leukaemia) or a propensity to carcinoma by determination of the
CC   of a CA gene, or by determination of CA gene expression in parti
CC   tissues. CA nucleic acids, proteins and antibodies are also usef
CC   therapeutic agents and in screening and evaluating drug candidat
CC   present sequence represents a specifically claimed murine CA nuc
CC   sequence of the invention. Note: The complete sequence data for
CC   patent did not form part of the printed specification, but was o
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 2942 BP; 702 A; 837 C; 780 G; 623 T; 0 U; 0 Other;

Query Match   1.6%;   Score 22;   DB 8;   Length 2942;
Best Local Similarity   100.0%;   Pred. No. 7.2;
Matches   22;   Conservative   0;   Mismatches   0;   Indels   0;

Qy   427   GCAGCCCATTATGAAGTTCATC 448
           ||||||||||||||||||||||
Db   552   GCAGCCCATTATGAAGTTCATC 573

RESULT 71
ADB72732
ID   ADB72732 standard; mRNA; 2942 BP.
XX
AC   ADB72732;
XX
DT   04-DEC-2003   (first entry)
XX
DE   Mouse Sept9 mRNA.
XX
KW   mouse; ss; cytostatic; gene therapy; vaccine; carcinoma; lymphoma
KW   cancer; neoplasm; adenocarcinoma; sarcoma.
XX
OS   Mus sp.
XX
PN   WO2003008583-A2.
XX
PD   30-JAN-2003.
XX
PF   26-DEC-2001; 2001WO-US051291.
XX
PR   02-MAR-2001; 2001US-00798586.
PR   23-OCT-2001; 2001US-00004113.

2001US-00052482.
2001US-00997722.
2001US-00034650.

S DISCOVERY.

ngelhard EK;

337/23.

nt nucleic acid, useful for treating carcinomas, lymphomas,
lasm, adenocarcinoma, or sarcomas.

ID NO 560; 2304pp; English.

relates to a novel recombinant nucleic acid comprising a
quence selected from any of the 660 sequences fully defined
ication. A polynucleotide of the invention has cytostatic
may have a use in gene therapy, or in a vaccine. The
ucleic acids and polypeptides are useful for treating
.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
present sequence represents a mouse mRNA of the invention.

BP; 702 A; 837 C; 780 G; 623 T; 0 U; 0 Other;
arity   1.6%; Score 22; DB 9; Pred. No. 7.2; Length 2942;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

CCATTATGAAGTTCATC 448
|||||||||||||||||
CCATTATGAAGTTCATC 573

lard; DNA; 2942 BP.

(first entry)

mRNA sequence.

ene therapy; vaccine; cancer; carcinoma-associated gene; CA;
smembrane; intracellular; ds.

2002WO-US038582.

2001US-00997722.

S DISCOVERY.

ngelhard EK;

03/48.

t nucleic acid comprising a nucleotide sequence of any of
associated (CA) genes, useful for screening for drug
diagnosing or treating carcinomas.

D NO 260; 983pp; English.

relates to a recombinant nucleic acid comprising a
quence selected from any of the fully defined carcinoma.
) genes from the 50 tables given in the specification. The
e secreted, transmembrane or intracellular proteins. The

CC   recombinant nucleic acids are useful for screening for drug cand
CC   for diagnosing or treating carcinomas. Sequences given in ADC852
CC   ADC85514 represent CA genes of the invention.
XX
SQ   Sequence 2942 BP; 702 A; 837 C; 780 G; 623 T; 0 U; 0 Other;

     Query Match
     Best Local Similarity   1.6%;   Score 22; DB 9; Length 2942;
                       100.0%; Pred. No. 7.2;
     Matches   22; Conservative   0; Mismatches   0; Indels   0; C

Qy   427 GCAGCCCATTATGAAGTTCATC 448
          ||||||||||||||||||||||
Db   552 GCAGCCCATTATGAAGTTCATC 573

RESULT 73
ADA02993
ID   ADA02993 standard; DNA; 50295 BP.
XX
AC   ADA02993;
XX
DT   06-NOV-2003   (first entry)
XX
DE   Mouse Sept9 carcinoma associated gene, SEQ ID NO:1511.
XX
KW   Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer;
KW   prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug scr
KW   gene; ds.
XX
OS   Mus sp.
XX
PN   WO200305746-A2.
XX
PD   17-JUL-2003.
XX
PF   26-DEC-2002; 2002WO-US041414.
XX
PR   26-DEC-2001; 2001US-00035832.
XX
PA   (SAGR-) SAGRES DISCOVERY.
XX
PI   Morris DW;
XX
DR   WPI; 2003-587068/55.
XX
PT   New recombinant nucleic acid encoding carcinoma associated protei
PT   useful for preparing compositions for treating carcinomas.
XX
PS   Claim 1; SEQ ID NO 1511; 245pp; English.
XX
CC   The invention relates to recombinant carcinoma associated (CA) nu
CC   acid sequences from mouse and human (ADA01482-ADA03094), and to
CC   recombinant carcinoma associated proteins (CAP) encoded by them.
CC   invention also encompasses expression vectors and host cells comp
CC   CA nucleic acid, a polypeptide (especially an antibody) that spec
CC   binds to the protein, and a biochip comprising CA nucleic acid or
CC   fragments thereof. The sequences of the invention were identified
CC   oncogenic retroviruses, which insert into the genome of the host c
CC   at random. Many of these do not carry transduced host oncogenes o
CC   pathogenic trans-acting viral genes, meaning that cancer incidence
CC   direct consequence of the effects of proviral integration into ho
CC   protooncogenes. The CA nucleic acid sequences can be used to diag
CC   carcinoma (especially breast cancer, prostate cancer, lymphoma or
CC   leukaemia) or a propensity to carcinoma by determination of the s
CC   of a CA gene, or by determination of CA gene expression in partic
CC   tissues. CA nucleic acids, proteins and antibodies are also useful
CC   therapeutic agents and in screening and evaluating drug candidate
CC   present sequence represents a specifically claimed murine CA nucle
CC   sequence of the invention. Note: The complete sequence data for t
CC   patent did not form part of the printed specification, but was obt
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX

XX
DT   01-JAN-2004 (first entry)
XX
DE   Mouse Sept19 genomic sequence.
XX
KW   Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated
KW   secreted; transmembrane; intracellular; ds.
XX
OS   Mus sp.
XX
PN   WO200304520-A2.
XX
PD   05-JUN-2003.
XX
PF   02-DEC-2002; 2002WO-US038582.
XX
PR   30-NOV-2001; 2001US-00997722.
XX
PA   (SAGR-) SAGRES DISCOVERY.
XX
PI   Morris DW, Engelhard EK;
XX
DR   WPI; 2003-513603/48.
XX
PT   New recombinant nucleic acid comprising a nucleotide sequence of
PT   the carcinoma-associated (CA) genes, useful for screening for dr
PT   candidates for diagnosing or treating carcinomas.
XX
PS   Claim 1; SEQ ID NO 259; 983pp; English.
XX
CC   The invention relates to a recombinant nucleic acid comprising a
CC   nucleotide sequence selected from any of the fully defined carci
CC   associated (CA) genes from the 50 tables given in the specificat
CC   CA proteins are secreted, transmembrane or intracellular protein
CC   recombinant nucleic acids are useful for screening for drug cand
CC   for diagnosing or treating carcinomas. Sequences given in ADC852
CC   ADC85514 represent CA genes of the invention.
XX
SQ   Sequence 50295 BP; 10891 A; 12938 C; 13611 G; 12237 T; 0 U; 618

Query Match   1.6%; Score 22; DB 9; Length 50295;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches   22; Conservative   0; Mismatches   0; Indels   0;

Qy   427 GCAGCCATTATGAAGTTCATC 448
          ||||||||||||||||||||
Db 31072 GCAGCCATTATGAAGTTCATC 31093

Search completed: April 8, 2004, 21:02:41
Job time : 662 secs

---

95 BP; 10891 A; 12938 C; 13611 G; 12237 T; 0 U; 618 Other;
larity   1.6%; Score 22; DB 8; Length 50295;
Conservative   0; Mismatches   0; Indels   0; Gaps   0;

GCCCATTATGAAGTTCATC 448
|||||||||||||||||||
GCCCATTATGAAGTTCATC 31093

ndard; DNA; 50295 BP.

(first entry)

gene.

ytostatic; gene therapy; vaccine; carcinoma; lymphomas;
lasm; adenocarcinoma; sarcoma; gene.

-A2.

2001WO-US051291.

2001US-00798586.
2001US-00004113.
2001US-00052482.
2001US-00997722.
2001US-00034650.

BS DISCOVERY.

Engelhard EK;

9337/23.

ant nucleic acid, useful for treating carcinomas, lymphomas,
plasm, adenocarcinoma, or sarcomas.

ID NO 559; 2304pp; English.

relates to a novel recombinant nucleic acid comprising a
equence selected from any of the 660 sequences fully defined
ication. A polynucleotide of the invention has cytostatic
may have a use in gene therapy, or in a vaccine. The
nucleic acids and polypeptides are useful for treating
e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
present sequence represents a mouse gene of the invention.

95 BP; 10891 A; 12938 C; 13611 G; 12237 T; 0 U; 618 Other;
arity 100.0%; Score 22; DB 9; Length 50295;
Conservative   0; Mismatches   0; Indels   0; Gaps   0;

CCCCATTATGAAGTTCATC 448
|||||||||||||||||||
CCCCATTATGAAGTTCATC 31093

ndard; DNA; 50295 BP.

Copyright GenCore version 5.1.6
(c) 1993 - 2004 Compugen Ltd.

in search, using sw model

ril 7, 2004, 17:54:48 ; Search time 20 Seconds
(without alignments)
1365.920 Million cell updates/sec

-09-245-198A-4
4
MSLLDFEISARRLPLPRSLG.........PWAHLKAAPFLTYFGLFQVH 284

IGO
pop 60.0 , Gapext 60.0

3366 seqs, 96191526 residues

ts satisfying chosen parameters: 283366

gth: 0
gth: 2000000000

isting first 100 summaries

IR_78:*
    pir1:*
    pir2:*
    pir3:*
    pir4:*

the number of results predicted by chance to have a
than or equal to the score of the result being printed,
ed by analysis of the total score distribution.

SUMMARIES

ch	Length	DB	ID	Description
.5	111	2	A85866	hypothetical prote
.2	733	2	S78376	photosystem I P700
.2	58	2	A58208	protamine I-1 - pa
.8	143	2	G84168	hypothetical prote
.8	197	2	E72374	hypothetical prote
.8	220	2	AG3547	bicyclomycin resis
.8	278	2	D83080	hypothetical prote
.8	339	2	C71132	hypothetical prote
.8	372	2	H70813	probable cysteine
.8	379	2	E64300	formate dehydrogen
.8	381	2	AH3041	conserved hypothet
.8	387	2	D84885	hypothetical prote
.8	397	2	D98244	hypothetical prote
.8	422	1	A60503	sperm-binding glyc
.8	443	2	T17220	hypothetical prote
.8	465	2	AC0347	probable membrane
.8	471	2	A75267	probable transport
.8	576	2	E64186	probable ATP-bindi
.8	586	2	A41125	gamma-glutamyltran
.5	45	2	D58208	protamine II-3 - p
.5	50	2	S22582	protamine 1 - Sagu
.5	58	2	S34045	protamine - North
.5	86	2	F87604	hypothetical Prote
.5	102	2	F87993	protein ZC34.3 [i
.5	115	2	PH1560	Ig heavy chain V r
.5	115	2	H83201	conserved hypothet
.5	118	1	IEEC5B	hypothetical prote
.5	118	2	AE1753	Orf51 [bacteriopha
.5	123	2	AH2707	conserved hypothet

No.		ID		Score		Description
30	2	T27519	125	2.5	?	hypothetica
31	2	I52290	131	2.5	?	interleukii
32	2	E30552	131	2.5	?	T-cell act:
33	2	T37116	146	2.5	?	probable ti
34	2	A71217	147	2.5	?	hypothetica
35	2	T08734	150	2.5	?	hypothetica
36	2	S31078	157	2.5	?	seed allerc
37	2	T02664	157	2.5	?	allergen -
38	2	A75567	157	2.5	?	conserved l
39	2	E75530	157	2.5	?	hypothetica
40	2	S59925	160	2.5	?	allergen RJ
41	1	DNEC17	161	2.5	?	outer membi
42	2	D90651	161	2.5	?	histone-lik
43	2	D85502	161	2.5	?	hypothetica
44	2	T24937	162	2.5	?	hypothetica
45	2	T31173	162	2.5	?	hypothetica
46	2	S76920	164	2.5	?	D2007.4 prc
47	2	S44789	170	2.5	?	transcripti
48	2	D87638	174	2.5	?	beta-lactog
49	1	LGGT	180	2.5	?	beta-lactoc
50	1	LGSH	180	2.5	?	insulin-lik
51	2	B60738	181	2.5	?	phospholipa
52	2	G85343	187	2.5	?	hypothetica
53	2	S72567	206	2.5	?	conserved h
54	2	AH0692	230	2.5	?	probable me
55	2	B64920	231	2.5	?	hypothetica
56	2	E90921	231	2.5	?	hypothetica
57	2	A85770	231	2.5	?	amino acid
58	2	C86665	231	2.5	?	ribonucleas
59	2	S60767	233	2.5	?	probable me
60	2	AB0273	233	2.5	?	H+-transpor
61	2	G85098	234	2.5	?	probable so
62	2	I64174	235	2.5	?	probable cy
63	2	A86387	236	2.5	?	hypothetica
64	2	H70866	238	2.5	?	conserved h
65	2	G83208	240	2.5	?	molybdopter
66	2	C64124	243	2.5	?	hypothetica
67	2	T25942	243	2.5	?	lymphotoxin
68	2	A46066	244	2.5	?	molybdopter
69	2	B32352	249	2.5	?	molybdopter
70	2	H90741	249	2.5	?	molybdopter
71	2	C85592	249	2.5	?	molybdopter
72	2	AI0602	252	2.5	?	IS21 family
73	2	AI3154	252	2.5	?	istB protei
74	2	H98132	254	2.5	?	molybdopter.
75	2	F82439	254	2.5	?	molybdopter
76	2	AB0182	255	2.5	?	hydroxypyru
77	2	AI0566	258	2.5	?	probable ac
78	2	G83101	258	2.5	?	lipopolysac
79	2	AG0007	260	2.5	?	flagellar b
80	2	F71801	262	2.5	?	flagellar b
81	2	A64718	262	2.5	?	probable bei
82	2	G75518	277	2.5	?	probable bel
83	2	A97396	293	2.5	?	hypothetica.
84	2	AB2614	293	2.5	?	hypothetica
85	2	E87280	300	2.5	?	rarD protei
86	2	S23780	303	2.5	?	nucleic aci
87	2	AE3466	307	2.5	?	maltose tra
88	2	D95932	308	2.5	?	probable su
89	2	D70328	310	2.5	?	histidine k
90	2	A31846	312	2.5	?	130K paracr
91	2	C75208	313	2.5	?	sugar trans
92	1	A46489	314	2.5	?	pan-epithel
93	2	C82561	316	2.5	?	drug tolera
94	2	AD3295	318	2.5	?	sodium/bile
95	2	C95927	319	2.5	?	probable su
96	2	B72604	321	2.5	?	hypothetica
97	2	C36718	325	2.5	?	pyruvate del
98	2	T18619	329	2.5	?	hypothetica
99	2	B90172	333	2.5	?	conserved h
100	2	A43577	343	2.5	?	regulatory

ALIGNMENTS

ein 23516 [imported] - Escherichia coli (strain O157:H7, substrain EDL9
ichia coli
001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
?66
unkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
ck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
33, 2001
sequence of enterohemorrhagic Escherichia coli O157:H7.
r: A85480; MUID:21074935; PMID:11206551
?66
nary
DNA
<STO>
s: GB:AE005174; NID:g12516604; PIDN:AAG57389.1; GSPDB:GN00145; UWGP:Z35
rce: strain O157:H7, substrain EDL933

larity   3.5%;   Score 10;   DB 2;   Length 111;
100.0%;   Pred. No. 0.074;
Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

ALACLGL 72
|||||||
ALACLGL 49

0 apoprotein A2 - Odontella sinensis chloroplast
plast Odontella sinensis
98 #sequence_revision 26-Feb-1998 #text_change 20-Jun-2000
?76
Stoebe, B.; Schaffran, I.; Kroth-Pancic, P.; Freier, U.
Rep. 13, 336-342, 1995
coplast Genome of a chlorophyll a+c containing Alga, Odontella sinensi
c: S78238
?76
nary; nucleic acid sequence not shown; translation not shown
DNA
<KOW>
s: EMBL:Z67753; NID:g1185127; PIDN:CAA91749.1; PID:g1185266
otide sequence was submitted to the EMBL Data Library, November 1995

last
otosystem I P700 apoprotein
oplast; electron transfer; membrane protein; membrane-associated comple

arity   3.2%;   Score 9;   DB 2;   Length 733;
100.0%;   Pred. No. 3.3;
onservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

LLACLG 71
||||||
LLACLG 341

inted turtle
ys picta (painted turtle)
6 #sequence_revision 08-Nov-1996 #text_change 07-May-1999
8
nsky, H.E.; Elsey, R.M.; Wright, C.L.; Rice, P.; Bell, J.E.; Sharp, D.
, 23547-23557, 1996
s of reptiles.
: A58208; MUID:96394458; PMID:8798564

A;Accession: A58208
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-58 <HUN>
C;Superfamily: sperm histone

Query Match   2.8%;   Score 8;   DB 2;   Length 58;
Best Local Similarity 100.0%;   Pred. No. 3.5;
Matches 8;   Conservative 0;   Mismatches 0;   Indels 0;

Qy   42 QRRRGRRG 49
Db   35 QRRRGRRG 42

RESULT 4
G84168
hypothetical protein Vng0080h [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Fe
C;Accession: G84168
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shu
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Ma
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt,
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: G84168
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-143 <STO>
A;Cross-references: GB:AE004437; NID:g10579733; PIDN:AAG18715.1; GSPD
C;Genetics:
A;Gene: VNG0080H

Query Match   2.8%;   Score 8;   DB 2;   Length 143;
Best Local Similarity 100.0%;   Pred. No. 7.6;
Matches 8;   Conservative 0;   Mismatches 0;   Indels 0;

Qy   61 LGLGLALA 68
Db   55 LGLGLALA 62

RESULT 5
E72374
hypothetical protein - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Ju
C;Accession: E72374
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.;
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacter
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: E72374
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-197 <ARN>
A;Cross-references: GB:AE001724; GB:AE000512; NID:g4980966; PIDN:AAD35
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM0469

Query Match   2.8%;   Score 8;   DB 2;   Length 197;
Best Local Similarity 100.0%;   Pred. No. 10;
Matches 8;   Conservative 0;   Mismatches 0;   Indels 0;

Qy   2 SLLDFEIS 9
Db   135 SLLDFEIS 142

A;Accession: C71132
A;Status: preliminary; nucleic acid sequence not shown; translation nc
A;Molecule type: DNA
A;Residues: 1-339 <KAW>
A;Cross-references: GB:AP000003; NID:g3236130; PIDN:BAA29917.1; PID:g3
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence re
C;Genetics:
A;Gene: PH0824
C;Superfamily: conserved hypothetical protein MTH900

Query Match   2.8%; Score 8; DB 2; Length 339;
Best Local Similarity 100.0%; Pred. No. 16;
Matches   8; Conservative 0; Mismatches 0; Indels 0; G

Qy   70 LGLLLAVV 77
         ||||||||
Db   111 LGLLLAVV 118

RESULT 9
H70813
probable cysteine synthase - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun
C;Accession: H70813
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Har
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamli
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squ
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barre
A;Title: Deciphering the biology of Mycobacterium tuberculosis from th
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: H70813
A;Status: preliminary; nucleic acid sequence not shown; translation no
A;Molecule type: DNA
A;Residues: 1-372 <COL>
A;Cross-references: GB:AL022004; GB:AL123456; NID:g3261550; PIDN:CAA17(
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: cysM3
C;Superfamily: threonine dehydratase

Query Match   2.8%; Score 8; DB 2; Length 372;
Best Local Similarity 100.0%; Pred. No. 17;
Matches   8; Conservative 0; Mismatches 0; Indels 0; G

Qy   61 LGLGLALA 68
         ||||||||
Db   96 LGLGLALA 103

RESULT 10
E64300
formate dehydrogenase (EC 1.2.1.2) beta chain - Methanococcus jannaschi
C;Species: Methanococcus jannaschii
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-
C;Accession: E64300
R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutt
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick,
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurs
Science 273, 1058-1073, 1996
A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smit
A;Title: Complete genome sequence of the methanogenic archaeon, Methano
A;Reference number: A64300; MUID:96337999; PMID:8688087
A;Accession: E64300
A;Status: preliminary; nucleic acid sequence not shown; translation not
A;Molecule type: DNA
A;Residues: 1-379 <BU>
A;Cross-references: GB:U67459; GB:L77117; NID:g2826236; PIDN:AAB97986.1
C;Genetics:
A;Map position: REV7250-6111
C;Superfamily: formate dehydrogenase chain B; ferredoxin 2[4Fe-4S] homo

---

tance protein [imported] - Brucella melitensis (strain 16M)
a melitensis
2 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
7
; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
man, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Sci. U.S.A. 99, 443-448, 2002
a sequence of the facultative intracellular pathogen Brucella melitens
; AD3252; PMID:11756688
ary
<KUR>
NA
. GB:AE008918; PIDN:AAL53546.1; PID:g17984455; GSPDB:GN00191
rce: strain 16M

rity   2.8%; Score 8; DB 2; Length 220;
100.0%; Pred. No. 11;
nservative 0; Mismatches 0; Indels 0; Gaps 0;

LAL 61
    |||
LAL 139

n PA4521 [imported] - Pseudomonas aeruginosa (strain PAO1)
nas aeruginosa
#sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

m, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
M.V.
, 2000
enome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A82950; MUID:20437337; PMID:10984043
ry
A
STO>
GB:AE004866; GB:AE004091; NID:g9950760; PIDN:AAG07909.1; GSPDB:GN001
ce: strain PAO1

rity   2.8%; Score 8; DB 2; Length 278;
100.0%; Pred. No. 13;
nservative 0; Mismatches 0; Indels 0; Gaps 0;

AVV 77
    |||
AVV 54

n PH0824 - Pyrococcus horikoshii
us horikoshii
#sequence_revision 14-Aug-1998 #text_change 20-Jun-2000

Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
nahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
1998
equence and gene organization of the genome of a hyper-thermophilic a
A71000; MUID:98344137; PMID:9679194

reductase
ferredoxin 2[4Fe-4S] homology <FER>

...larity 2.8%; Score 8; DB 2; Length 379;
Conservative 100.0%; Pred. No. 17; Mismatches 0; Indels 0; Gaps 0;

```
GVLAL 224
||||||
GVLAL 42
```

...ical protein Atu3948 [imported] - Agrobacterium tumefaciens (strain C5
...cterium tumefaciens
...02 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
...1
...bal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
..., W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
...D., P.; Zhang, S.
...-2323, 2001
...; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
...ne of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
...C: AB2577; MUID:21608550; PMID:11743193
...ary
...DNA
...<KUR>
...3: GB:AE008689; PIDN:AAL44750.1; PID:g17742385; GSPDB:GN00187
...rce: strain C58 (Dupont)

...inear chromosome

...arity 2.8%; Score 8; DB 2; Length 381;
...100.0%; Pred. No. 17;
...onservative 0; Mismatches 0; Indels 0; Gaps 0;

```
PGSS 258
|||
PGSS 332
```

...in At2g45000 [imported] - Arabidopsis thaliana
...psis thaliana (mouse-ear cress)
...1 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
...5
...; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
...t, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
...W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
...8, 1999
...and analysis of chromosome 2 of the plant Arabidopsis thaliana.
...: A84420; MUID:20083487; PMID:10617197
...5
...ary
...DNA
...<STO>
...: GB:AE002093; NID:g4895250; PIDN:AAD32835.1; GSPDB:GN00139

...arity 2.8%; Score 8; DB 2; Length 387;
...100.0%; Pred. No. 18;
...onservative 0; Mismatches 0; Indels 0; Gaps 0;

```
AEED 100
|||
AEED 377
```

RESULT 13
D98244
hypothetical protein AGR_L_1808 [imported] - Agrobacterium tumefacien
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-No
C;Accession: D98244
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qu
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappa
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agen
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: D98244
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-397 <KUR>
A;Cross-references: GB:AE007870; PIDN:AAK89478.1; PID:g15159347; GSPD
C;Genetics:
A;Gene: AGR_L_1808
A;Map position: linear chromosome

Query Match 2.8%; Score 8; DB 2; Length 397;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0;

```
Qy  251 LALRPGSS 258
          ||||||||
Db  341 LALRPGSS 348
```

RESULT 14
A60503
sperm-binding glycoprotein ZP3 precursor - golden hamster
N;Alternate names: sperm receptor; zona pellucida glycoprotein ZP3
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Se
C;Accession: A60503
R;Kinloch, R.A.; Ruiz-Seiler, B.; Wassarman, P.M.
Dev. Biol. 142, 414-421, 1990
A;Title: Genomic organization and polypeptide primary structure of zo
A;Reference number: A60503; MUID:91078540; PMID:2257975
A;Accession: A60503
A;Molecule type: DNA
A;Residues: 1-422 <KIN>
A;Cross-references: GB:M63629
A;Note: the authors translated the codon CAA for residue 251 as Glu, a
C;Comment: This sulfated glycoprotein in the zona pellucida of the oo
C;Superfamily: sperm-binding glycoprotein ZP3; ZP domain homology
C;Keywords: glycoprotein; oocyte
F;45-300/Domain: ZP domain homology <ZPH>

Query Match 2.8%; Score 8; DB 1; Length 422;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0;

```
Qy  59 LAIGLGLA 66
        ||||||||
Db  386 LAIGLGLA 393
```

RESULT 15
T17220
hypothetical protein DKFZp5660011.1 - human
C;Species: Homo sapiens (man)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Jun
C;Accession: T17220
R;Blum, H.; Bauerbachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A;Reference number: Z18725
A;Accession: T17220
A;Status: preliminary
A;Molecule type: mRNA

```
<BLU>
: EMBL:AL117414
rce: fetal kidney; clone DKFZp5560011

11.1
ma-glutamyltransferase

arity    2.8%;   Score 8;  DB 2;   Length 443;
100.0%;  Pred. No. 20;
onservative   0;  Mismatches    0;  Indels    0;  Gaps    0;

LALA 68
     |||
LALA 21
```

```
protein yegB [imported] - Yersinia pestis (strain CO92)
a pestis
l #sequence_revision 02-Nov-2001 #text_change 17-May-2002

en, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
, 2001
uence of Yersinia pestis, the causative agent of plague.
; AB0001; MUID:21470413; PMID:11586360
ry
VA
:KUR>
: GB:AL590842; PIDN:CAC92102.1; PID:g15980820; GSPDB:GN00175
```

```
idrug-efflux transporter

urity    2.8%;   Score 8;  DB 2;   Length 465;
100.0%;  Pred. No. 21;
onservative   0;  Mismatches    0;  Indels    0;  Gaps    0;

VSL 79
    |||
VSL 342
```

```
protein - Deinococcus radiodurans (strain R1)
cus radiodurans
#sequence_revision 03-Dec-1999 #text_change 31-Mar-2000

J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
athevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
nter, J.C.; Fraser, C.M.
577, 1999
uence of the radioresistant bacterium Deinococcus radiodurans R1.
A75250; MUID:20036896; PMID:10567266
ry
A
WHI>
: GB:AE002079; GB:AE000513; NID:g6460315; PIDN:AAF12043.1; PID:g646032
ce: strain R1
```

```
rity    2.8%;   Score 8;  DB 2;   Length 471;
100.0%;  Pred. No. 21;
nservative   0;  Mismatches    0;  Indels    0;  Gaps    0;

LAL 67
```

```
Db       366  ALGLGLAL  373
                  ||||||||
```

```
RESULT 18
B64186
probable ATP-binding transport protein HI1156 - Haemophilus influenzae
C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 02-Feb
C;Accession: E64186
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness,
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley
, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Ge
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smi
A;Title: Whole-genome random sequencing and assembly of Haemophilus in
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: E64186
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-576 <TIGR>
A;Cross-references: GB:U32795; GB:L42023; NID:g1574708; PIDN:AAC22811.
C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding c
C;Keywords: ATP; nucleotide binding; P-loop
F;355-550/Domain: ATP-binding cassette homology <ABC>
F;372-379/Region: nucleotide-binding motif A (P-loop)

Query Match            2.8%;   Score 8;  DB 2;  Length 576;
Best Local Similarity  100.0%;  Pred. No. 25;
Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  G

Qy      58  PLALGLGL  65
                ||||||||
Db     159  PLALGLGL  166
```

```
RESULT 19
A41125
gamma-glutamyltransferase (EC 2.3.2.2) related protein - human
N;Alternate names: gamma-glutamyltransferase-like activity 1; GGT-REL
C;Species: Homo sapiens (man)
C;Date: 27-Mar-1992 #sequence_revision 27-Mar-1992 #text_change 18-Jun
C;Accession: A41125
R;Heisterkamp, N.; Rajpert-De Meyts, E.; Uribe, L.; Forman, H.J.; Grof
Proc. Natl. Acad. Sci. U.S.A. 88, 6303-6307, 1991
A;Title: Identification of a human gamma-glutamyl cleaving enzyme rela
A;Reference number: A41125; MUID:91296809; PMID:1676842
A;Accession: A41125
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-586 <HEI>
A;Cross-references: GB:M64099; NID:g183141; PIDN:AAA58503.1; PID:g1831
C;Genetics:
A;Gene: GDB:GGT1A1; GGT-REL
A;Cross-references: GDB:134033
C;Superfamily: gamma-glutamyltransferase
C;Keywords: aminoacyltransferase; glycoprotein; transmembrane protein

Query Match            2.8%;   Score 8;  DB 2;  Length 586;
Best Local Similarity  100.0%;  Pred. No. 25;
Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Ge

Qy      61  LGLGLALA  68
                ||||||||
Db      14  LGLGLALA  21
```

```
RESULT 20
D58208
protamine II-3 - painted turtle
C;Species: Chrysemys picta (painted turtle)
C;Date: 08-Nov-1996 #sequence_revision 08-Nov-1996 #text_change 07-May-
C;Accession: D58208
```

insky, H.E.; Elsey, R.M.; Wright, C.L.; Rice, P.; Bell, J.E.; Sharp, D.
1, 23547-23557, 1996
es of reptiles.
r: A58208; MUID:96394458; PMID:8798564
08
nary
protein
<HUN>
erm histone

larity   2.5%; Score 7; DB 2; Length 45;
100.0%; Pred. No. 26;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GRRG 49
|||
GRRG 30

RESULT 23
F87604
hypothetical protein CC2870 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Ap
C;Accession: F87604
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen,
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter,
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: F87604
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-86 <STO>
A;Cross-references: GB:AE005673; NID:g13424486; PIDN:AAK24834.1; GSPI
C;Genetics:
A;Gene: CC2870

Query Match   2.5%;   Score 7; DB 2; Length 86;
Best Local Similarity   100.0%; Pred. No. 45;
Matches   7; Conservative   0; Mismatches   0; Indels   0;

Qy   59 LALGLGL 65
         |||||||
Db   26 LALGLGL 32

uinus imperator
us imperator
93 #sequence_revision 12-Apr-1996 #text_change 21-Jul-2000
82
iva, R.
, 19, 5786, 1991
e 1 gene sequence from the primate Saguinus imperator isolated with PCR
r: S22582; MUID:92051332; PMID:1840669
82
DNA
<QUE>
s: EMBL:X61678; NID:gS8405; PIDN:CAA43853.1; PID:g4494091
rs translated the codon TAC for residue 43 as Thr

erm histone
osomal protein; DNA binding; nucleus; spermatogenesis

larity   2.5%; Score 7; DB 2; Length 50;
100.0%; Pred. No. 28;
Conservative   0; Mismatches   0; Indels   0; Gaps   0;

RGRR 48
||||
RGRR 24

RESULT 24
F87993
protein ZC334.3 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-Ma
C;Accession: F87993
R;anonymous; The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for i
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.a
A;Note: published errata appeared in Science 283, 35, 1999; Science 2
A;Accession: F87993
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-102 <STO>
A;Cross-references: GB:chr_I; PIDN:CAB04964.1; PID:g3881432; GSPDB:GN
A;Note: predicted using Genefinder
C;Genetics:
A;Gene: ZC334.3
A;Map position: 1

Query Match   2.5%;   Score 7; DB 2; Length 102;
Best Local Similarity   100.0%; Pred. No. 52;
Matches   7; Conservative   0; Mismatches   0; Indels   0;

Qy   55 LLVPLAL 61
         |||||||
Db   33 LLVPLAL 39

American opossum
nis virginiana; Didelphis marsupialis virginiana (North American opossu
95 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
15
Nishikawa, S.; Connor, W.; Dixon, G.H.
215, 63-72, 1993
ization of a marsupial sperm protamine gene and its transcripts from b
C: S34045; MUID:93345500; PMID:8344286
15
nary
DNA
<WIN>
s: EMBL:X74044; NID:g407062; PIDN:CAA52193.1; PID:g407063
erm histone
nding; nucleus

arity   2.5%; Score 7; DB 2; Length 58;
100.0%; Pred. No. 32;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

RRG 49
|||
RRG 41

RESULT 25
PHI560
Ig heavy chain V region (clone VH32) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 05-Aug-1994 #sequence_revision 05-Aug-1994 #text_change 21-Jan
C;Accession: PHI560
R;Rassenti, L.Z.; Kipps, T.J.
J. Exp. Med. 177, 1039-1046, 1993
A;Title: Lack of extensive mutations in the VH5 genes used in common E
A;Reference number: PHI557; MUID:93210459; PMID:7681468
A;Accession: PHI560

Qy   23 DGGAVRQ 29
          |||||||
Db   107 DGGAVRQ 113

RESULT 28
AE1753
Orf51 [bacteriophage bIL285] homolog lin2570 [imported] - Listeria inn
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov
C;Accession: AE1753
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Be
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entia
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.;
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.;
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AE1753
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-118 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC97797.1; PID:g16415092; GSPDB
A;Experimental source: strain Clipl1262
C;Genetics:
A;Gene: lin2570

Query Match   2.5%; Score 7; DB 2; Length 118;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; G

Qy   2 SLLDFEI 8
        |||||||
Db   61 SLLDFEI 67

RESULT 29
AH2707
conserved hypothetical protein Atu1065 [imported] - Agrobacterium tume
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov
C;Accession: AH2707
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.
erage, G.; Gillet, W.; Grant, S.; Guenthner, D.; Kutyavin, T.; Levy, R
.; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tume
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AH2707
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-123 <KUR>
A;Cross-references: GB:AE008688; PIDN:AAL42078.1; PID:g17739458; GSPDB
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu1065
A;Map position: circular chromosome

Query Match   2.5%; Score 7; DB 2; Length 123;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Ga

Qy   12 RLPLPRS 18
        |||||||
Db   50 RLPLPRS 56

RESULT 30
T27519
hypothetical protein ZC334.3 - Caenorhabditis elegans

NA
<RAS>
nnoglobulin v region; immunoglobulin homology
etramer; immunoglobulin
mmunoglobulin homology <IMM>

arity   2.5%; Score 7; DB 2; Length 115;
onservative 0; Mismatches 0; Indels 0; Gaps 0;

AV 76
||
AV 11

ical protein PA3557 [imported] - Pseudomonas aeruginosa (strain PAO1)
onas aeruginosa
0 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
L
am, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
M.V.
, 2000
genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A82950; MUID:20437337; PMID:10984043

ary
NA
<STO>
GB:AE004776; GB:AE004091; NID:g9949701; PIDN:AAG06945.1; GSPDB:GN001
ce: strain PAO1

ary   2.5%; Score 7; DB 2; Length 115;
onservative 0; Mismatches 0; Indels 0; Gaps 0;

GL 65
||
GL 55

n, 12K - Escherichia coli insertion sequence IS5
chia coli
#sequence_revision 18-Dec-1981 #text_change 10-Sep-1999
; C03582; A04466
, M.
981
tide sequence of IS5 from Escherichia coli.
A91483; MUID:82028653; PMID:6269959

A
SCH>
Bree, M.P.
981
tide sequence and protein-coding capability of the transposable eleme
A03582; MUID:82028652; PMID:6269958
A
ENG>

nsertion sequence IS5
erichia coli hypothetical protein o263

rity   2.5%; Score 7; DB 1; Length 118;
onservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches   7; Conservative   0; Mismatches   0; Indels   0;

Qy   65 LALACLG 71
           |||||||
Db   9 LALACLG 15

RESULT 33
T37116
probable transposase, truncated [imported] - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 15-Se
C;Accession: T37116
R;Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B
submitted to the EMBL Data Library, August 1999
A;Reference number: Z21588
A;Accession: T37116
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-146 <SAU>
A;Cross-references: EMBL:AL109950; PIDN:CAB52967.1; GSPDB:GN00070; SC(
A;Experimental source: strain A3(2)
C;Genetics:
C;Gene: SCOEDB:SCU4.33c
C;Superfamily: Synechocystis transposase sll710

Query Match   2.5%; Score 7; DB 2;   Length 146;
Best Local Similarity 100.0%; Pred. No. 70;
Matches   7; Conservative   0; Mismatches   0; Indels   0; (

Qy   70 LGLLLAV 76
           |||||||
Db   102 LGLLLAV 108

RESULT 34
A71217
hypothetical protein PH2001 - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 20-Jur
C;Accession: A71217
R;Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; }
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; H
DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hy
A;Reference number: A71000; MUID:98344137; PMID:9679194
A;Accession: A71217
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-147 <KAW>
A;Cross-references: GB:AP000007; GB:AP000001; NID:g3236134; NID:g32361
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence re
A;Note: this sequence is split into two separate translations in GenBa
C;Genetics:
A;Gene: PH2001
C;Superfamily: Pyrococcus horikoshii hypothetical protein PH2001

Query Match   2.5%; Score 7; DB 2;   Length 147;
Best Local Similarity 100.0%; Pred. No. 71;
Matches   7; Conservative   0; Mismatches   0; Indels   0; G

Qy   67 LACLGLL 73
           |||||||
Db   45 LACLGLL 51

RESULT 35
T08734
hypothetical protein DKFZp566F0546.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 13-Aug
C;Accession: T08734

---

abditis elegans
9 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
9

EMBL Data Library, November 1996
: Z20381
19

ary; translated from GB/EMBL/DDBJ
DNA
<WIL>
: EMBL:Z82082; PIDN:CAB04964.2; GSPDB:GN00019; CESP:ZC334.3
rce: clone ZC334

4.3

larity   2.5%; Score 7; DB 2;   Length 125;
onservative   100.0%; Pred. No. 62;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

PLAL 61
   |||
PLAL 39

rat
norvegicus (Norway rat)
6 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999
0
uet, E.N.
Res. Commun. 197, 612-618, 1993
of rat interleukin-13 (IL-13) cDNA and analysis of IL-13 gene expressio
: I52290; MUID:94092138; PMID:7916615
0

ary; translated from GB/EMBL/DDBJ
RNA
<RES>
: GB:L26913; NID:g438875; PIDN:AAA16478.1; PID:g438876

erleukin-13

arity   2.5%; Score 7; DB 2;   Length 131;
100.0%; Pred. No. 64;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

CLG 71
|||
CLG 15

protein P600 precursor - mouse
culus (house mouse)
9 #sequence_revision 28-Aug-1989 #text_change 16-Jul-1999
2
awski, S.M.; Mosmann, T.R.; Zurawski, G.
79-687, 1989
of small inducible proteins secreted by leukocytes are members of a ne
ation processes.
: A30552; MUID:89093958; PMID:2521353
RNA
<BRO>
: GB:M23504; NID:g533246; PIDN:AAA40149.1; PID:g533247
erleukin-13

arity   2.5%; Score 7; DB 2;   Length 131;
100.0%; Pred. No. 64;

; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
rotein Sequence Database, May 1999
; Z16474
4
RNA
<OTT>
; EMBL.AL050075
rce: fetal kidney; clone DKFZp566F0546
546.1

arity   2.5%;   Score 7; DB 2; Length 150;
         100.0%; Pred. No. 72;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

LGP 240
        |||
LGP 38

- rice
ativa (rice)
3 #sequence_revision 30-Sep-1993 #text_change 20-Jun-2000
3
L, H.; Yamada, T.; Tanaka, K.; Takeuchi, S.; Nakamura, R.; Matsuda, T.
L, 239-248, 1993
cture and expression of rice seed allergenic proteins belonging to the
; S31078; MUID:9144699; PMID:7678765
3
RNA
:ADA>
EMBL:D11430; NID:g218196; PIDN:BAA01996.1; PID:g218197
t alpha-amylase inhibitor

rity   2.5%;   Score 7; DB 2; Length 157;
        100.0%; Pred. No. 75;
nservative   0; Mismatches   0; Indels   0; Gaps   0;

VS 78
||
VS 17

tiva (rice)
#sequence_revision 24-Mar-1999 #text_change 16-Jul-1999

J.H.; Eun, M.Y.
BL Data Library, January 1998
eotide sequence of rice allergenic protein.
Z14691
ry; translated from GB/EMBL/DDBJ
NA
YUN>
EMBL:AF042200; NID:g2827315; NID:g2827315; PIDN:AAB99797.1; PID:g2827316
ce: strain Nipponbare
t alpha-amylase inhibitor

rity   2.5%;   Score 7; DB 2; Length 157;
        100.0%; Pred. No. 75;
nservative   0; Mismatches   0; Indels   0; Gaps   0;

VS 78
||
VS 17

RESULT 38
A75567
conserved hypothetical protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar
C;Accession: A75567
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.
, S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus r
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: A75567
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-157 <WHI>
A;Cross-references: GB:AE001867; GB:AE000513; NID:g6457693; PIDN:AAF09
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR0033
A;Map position: 1

Query Match   2.5%;   Score 7; DB 2; Length 157;
Best Local Similarity   100.0%; Pred. No. 75;
Matches   7; Conservative   0; Mismatches   0; Indels   0; G

Qy   261 IRTLPWA 267
               |||||||
Db   22 IRTLPWA 28

RESULT 39
E75530
hypothetical protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar
C;Accession: E75530
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.
, S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus r
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: E75530
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-157 <WHI>
A;Cross-references: GB:AE001895; GB:AE000513; NID:g6458024; PIDN:AAF09
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR0352
A;Map position: 1

Query Match   2.5%;   Score 7; DB 2; Length 157;
Best Local Similarity   100.0%; Pred. No. 75;
Matches   7; Conservative   0; Mismatches   0; Indels   0; G

Qy   43 RRRGRRG 49
               |||||||
Db   144 RRRGRRG 150

RESULT 40
S59925
allergen RA5B precursor - rice
C;Species: Oryza sativa (rice)
C;Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 20-Jun-
C;Accession: S59925
R;Alvarez, A.M.; Adachi, T.; Nakase, M.; Aoki, N.; Nakamura, R.; Matsud
Biochim. Biophys. Acta 1251, 201-204, 1995
A;Title: Classification of rice allergenic protein cDNAs belonging to t
A;Reference number: S59922; MUID:95399441; PMID:7669811
A;Accession: S59925

Best Local Similarity 100.0%; Pred. No. 77;
Matches 7; Conservative 0; Mismatches 0; Indels 0;

```
Qy  62  GLGLALA  68
            |||||||
Db   9  GLGLALA  15
```

RESULT 42
D90651
histone-like protein HlpA [imported] - Escherichia coli (strain O157:
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Au
C;Accession: D90651
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yoko
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shina
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia co
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: D90651
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-161 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB33603.1; PID:g13359636; GSPD
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: EC80180
C;Superfamily: DNA-binding 17K protein

Query Match 2.5%; Score 7; DB 2; Length 161;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 7; Conservative 0; Mismatches 0; Indels 0;

```
Qy  62  GLGLALA  68
            |||||||
Db   9  GLGLALA  15
```

RESULT 43
D85502
hypothetical protein hlpA [imported] - Escherichia coli (strain O157:
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep
C;Accession: D85502
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.;
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potam
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: D85502
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-161 <STO>
A;Cross-references: GB:AE005174; NID:g12512906; PIDN:AAG54480.1; GSPD
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: hlpA
C;Superfamily: DNA-binding 17K protein

Query Match 2.5%; Score 7; DB 2; Length 161;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 7; Conservative 0; Mismatches 0; Indels 0;

```
Qy  62  GLGLALA  68
            |||||||
Db   9  GLGLALA  15
```

RESULT 44
T24937
hypothetical protein W03C9.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct

---

ary
RNA
<ALV>
: EMBL:D42142; NID:g1398917; PIDN:BAA07713.1; PID:g1398918
at alpha-amylase inhibitor

arity 100.0%; Pred. No. 76;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
VVS  78
|||
VVS  17
```

tein hlpA precursor - Escherichia coli (strain K-12)
DNA-binding 17K protein; histone-like protein hlp
chia coli
9 #sequence_revision 30-Jun-1989 #text_change 01-Mar-2002
4; A38063; S13728; B64742; I54944; S20426
e, K.
1988
nd sequencing of the gene for the DNA-binding 17K protein of Escherich
: JT0304; MUID:88329735; PMID:2843433
4
DNA
<HO1>
: GB:M21118; NID:g147821; PIDN:AAA24630.1; PID:g147822
3
protein
<HO2>
rce: strain B
etharam, S.
334-344, 1991
nd nucleotide sequence of the firA gene and the firA200(Ts) allele fro
8
: S13728; MUID:91100302; PMID:1987124

ary
DNA
<DIC>
: EMBL:X54797; NID:g41468; PIDN:CAA38567.1; PID:g41469
rce: strain K-12, substrain MG1655
Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
lau, B.; Shao, Y.
1462, 1997
ete genome sequence of Escherichia coli K-12.
: A64720; MUID:97426617; PMID:9278503
2
acid sequence not shown; translation not shown
DNA
<BLAT>
: GB:AE000127; GB:U00096; NID:g1786370; PIDN:AAC73289.1; PID:g1786375
rce: strain K-12, substrain MG1655
i, P.; Vaara, M.
1223-1229, 1991
gene of Yersinia enterocolitica: cloning, sequencing, expression, and
: I54944; MUID:91123198; PMID:1991717
ed from GB/EMBL/DDBJ
L', 16-148, 'E', 150-152, 'I', 154-161 <RES>
: EMBL:X75465; NID:g432661; PIDN:CAA53207.1; PID:g432662
A protein has been believed to be a histone-like constituent of bacter

-binding 17K protein
ne protein
nal sequence #status predicted <SIG>
outer membrane protein hlpA #status predicted <MAT>

2.5%; Score 7; DB 1; Length 161;

7; T26123

MBL Data Library, July 1995
; Z19957
7
ary; translated from GB/EMBL/DDBJ
NA
<WIL>
: EMBL:Z50015; PIDN:CAA90314.1; GSPDB:GN00020; CESP:W03C9.4
rce: clone T15G9
scough. R.
MBL Data Library, October 1995
: Z20155
3
ary; translated from GB/EMBL/DDBJ
NA
<WI2>
: EMBL:Z66516; PIDN:CAA91361.1; GSPDB:GN00020; CESP:W03C9.4
rce: clone W03C9
.4
4/1; 120/3; 128/1

arity   2.5%;   Score 7; DB 2; Length 162;
          100.0%; Pred. No. 77;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

ASS 237
|||
ASS 118

in 424 - Sphingomonas aromaticivorans plasmid pNL1
monas aromaticivorans
0 #sequence_revision 11-Jan-2000 #text_change 09-Jun-2000
3
llwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.; G
MBL Data Library, July 1998
plete sequence of a 184 kb catabolic plasmid from Sphingomonas aromati
Z20992
3
ary; translated from GB/EMBL/DDBJ
NA
:ROM>
: EMBL:AF079317; NID:g3378261; PID:g3378314; PIDN:AAD03897.1
pNL1

.ngomonas aromaticivorans hypothetical protein 424

arity   2.5%;   Score 7; DB 2; Length 162;
          100.0%; Pred. No. 77;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

AA 144
||
AA 34

.n - Synechocystis sp. (strain PCC 6803)
ystis sp.
, #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
)
S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
., 1996
analysis of the genome of the unicellular cyanobacterium Synechocystis

---

s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S76920
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-164 <KAN>
A;Cross-references: EMBL:D90917; GB:AB001339; NID:g1653836; PIDN:BAA1e
A;Note: the nucleotide sequence was submitted to the EMBL Data Library

Query Match   2.5%;   Score 7; DB 2; Length 164;
Best Local Similarity   100.0%; Pred. No. 78;
Matches   7; Conservative   0; Mismatches   0; Indels   0; C

Qy   104 SELNPQT 110
             |||||||
Db   37 SELNPQT 43

RESULT 47
S44789
D2007.4 protein - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 14-Sep-1994 #sequence_revision 12-May-1995 #text_change 23-Mar
C;Accession: S44789
R;Favello, A.D.
submitted to the EMBL Data Library, May 1993
A;Description: Sequence of the C. elegans cosmid D2007.
A;Reference number: S44619
A;Accession: S44789
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-170 <FAV>
A;Cross-references: EMBL:L16560; NID:g289666; PID:g289670
C;Genetics:
A;Introns: 43/2; 121/3

Query Match   2.5%;   Score 7; DB 2; Length 170;
Best Local Similarity   100.0%; Pred. No. 80;
Matches   7; Conservative   0; Mismatches   0; Indels   0; G

Qy   221 VLALRCL 227
             |||||||
Db   94 VLALRCL 100

RESULT 48
D87638
transcription regulator, GntR family [imported] - Caulobacter crescent
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr
C;Accession: D87638
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.B.; Eisen,
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.;
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter,
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: D87638
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-174 <STO>
A;Cross-references: GB:AE005673; NID:g13424808; PIDN:AAK25104.1; GSPDB
C;Genetics:
A;Gene: CC3142

Query Match   2.5%;   Score 7; DB 2; Length 174;
Best Local Similarity   100.0%; Pred. No. 82;
Matches   7; Conservative   0; Mismatches   0; Indels   0; G

Qy   138 ARRAIAA 144
             |||||||
Db   151 ARRAIAA 157

n precursor - goat
aegagrus hircus (domestic goat)
85 #sequence_revision 12-Apr-1996 #text_change 22-Jun-1999
20; S14507; S42800; S42801
unitzer, G.; Schrank, B.; Stangl, A.
  Physiol. Chem. 360, 1595-1604, 1979
o acid sequence of goat beta-lactoglobulin.
r: A91682; MUID:80070611; PMID:511095
20
protein
0 <PRE>
11, A.; Sanchez, A.
EMBL Data Library, March 1991
r: S14507
07
mRNA
<FOL>
s: EMBL:X58471; NID:g967; PIDN:CAA41385.1; PID:g968

EMBL Data Library, January 1993
r: S42800
00
mRNA
<KIM>
s: EMBL:Z19569; NID:g437751; PIDN:CAA79623.1; PID:g437752
01
mRNA
<KI2>
s: EMBL:Z19570; NID:g437753; PIDN:CAA79624.1; PID:g437754
physiological conditions beta-lactoglobulin exists as an equilibrium mi
pocalin; lipocalin homology

gnal sequence #status predicted <SIG>
beta-lactoglobulin #status predicted <MAT>
lipocalin homology <LIP>
Disulfide bonds: #status predicted

larity   2.5%;   Score 7;   DB 1;   Length 180;
Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

ALAC 69
|||
LAC 14

l precursor - sheep
: beta-lactoglobulin A; beta-lactoglobulin B; beta-lactoglobulin C; bet
cientalis aries, Ovis ammon aries (domestic sheep)
85 #sequence_revision 19-Apr-1996 #text_change 22-Jun-1999
19; JQ0748; A30011; B30011; S02136; A25136; A03221; S04955
aghan, M.; Simons, J.P.; Clark, A.J.
1990
cisation of the alleles encoding ovine beta-lactoglobulins A and B.
c: JQ0748; MUID:91007276; PMID:1976573
19
DNA
<ALIB>
s: GB:M32232
irce: beta-lactoglobulin B
18
DNA
'Y', 39-180 <ALIA>
s: GB:M32232
irce: beta-lactoglobulin A
A.J.
415-426, 1988
cization of the gene encoding ovine beta-lactoglobulin. Similarity to b

A;Reference number: A92942; MUID:88172489; PMID:3351935
A;Accession: A30011
A;Molecule type: DNA
A;Residues: 1-180 <ALI1>
A;Cross-references: GB:X14971
A;Experimental source: beta-lactoglobulin I
A;Accession: B30011
A;Molecule type: DNA
A;Residues: 1-37, 'Y',39-102,'N',104-180 <ALI2>
A;Cross-references: GB:X07009
A;Experimental source: beta-lactoglobulin II
R;Harris, S.; Ali, S.; Anderson, S.; Archibald, A.L.; Clark, A.J.
Nucleic Acids Res. 16, 10379-10380, 1988
A;Title: Complete nucleotide sequence of the genomic ovine beta-lacto
A;Reference number: S02136; MUID:89057492; PMID:3194215
A;Accession: S02136
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-180 <HAR>
A;Cross-references: EMBL:X12817; NID:g1313; PIDN:CAA31305.1; PID:g131
R;Gaye, P.; Hue-Delahaie, D.; Mercier, J.C.; Soulier, S.; Vilotte, J.
Biochimie 68, 1097-1107, 1986
A;Title: Ovine beta-lactoglobulin messenger RNA: nucleotide sequence
A;Reference number: A25136; MUID:87049827; PMID:3096387
A;Accession: A25136
A;Molecule type: mRNA
A;Residues: 1-180 <GAY>
A;Cross-references: GB:X04520; NID:g1315; PIDN:CAA28204.1; PID:g1316
R;Preaux, G.; Braunitzer, G.; Kolde, H.J.
Arch. Int. Physiol. Biochim. 88, B45-B46, 1980
A;Title: Primary structure of ovine beta-lactoglobulin.
A;Reference number: A03221; MUID:80219294; PMID:6155855
A;Accession: A03221
A;Molecule type: protein
A;Residues: 19-37,'Y',39-180 <PRE>
R;Erhardt, G.; Godovac-Zimmermann, J.; Conti, A.
Biol. Chem. Hoppe-Seyler 370, 757-762, 1989
A;Title: Isolation and complete primary sequence of a new ovine wild-
A;Reference number: S04955; MUID:89374823; PMID:2775495
A;Accession: S04955
A;Molecule type: protein
A;Residues: 19-37,'Y',39-165,'Q',167-180 <ERH>
A;Experimental source: beta-lactoglobulin C
C;Comment: This protein is the major milk whey protein of ruminants a
C;Comment: Under physiological conditions beta-lactoglobulin exists a
C;Genetics:
C;Gene: BLG
A;Introns: 32/3; 79/2; 104/1; 141/1; 176/1
C;Superfamily: lipocalin; lipocalin homology
C;Keywords: milk; polymorphism
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-180/Product: beta-lactoglobulin #status experimental <MAT>
F;28-178/Domain: lipocalin homology <LIP>
F;84-178,124-137/Disulfide bonds: #status predicted

Query Match   2.5%;   Score 7;   DB 1;   Length 180;
Best Local Similarity   100.0%;   Pred. No. 84;
Matches   7;   Conservative   0;   Mismatches   0;   Indels   0;   (

Qy   63 LGLALAC 69
        |||||||
Db   8 LGLALAC 14

RESULT 51
B60738
insulin-like growth factor II precursor - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 28-Apr-1993 #sequence_revision 30-Sep-1993 #text_change 13-No
C;Accession: S12614; B60738
R;Catchpole, I.R.; Engstrom, W.
Nucleic Acids Res. 18, 6430, 1990
A;Title: Nucleotide sequence of a porcine insulin-like growth factor I

C;Genetics:
A;Map position: 3
A;Introns: 42/2; 88/3; 126/3; 176/3
A;Note: C35D10.8
C;Superfamily: Caenorhabditis elegans hypothetical protein C35D10.8

Query Match   2.5%;   Score 7;   DB 2;   Length 206;
Best Local Similarity   100.0%;   Pred. No. 94;
Matches   7;   Conservative   0;   Mismatches   0;   Indels   0;   G

Qy   136 TRARRAI 142
           |||||||
Db   98 TRARRAI 104

RESULT 54
AH0692
conserved hypothetical protein ydgQ [imported] - Salmonella enterica s
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov
C;Accession: AH0692
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; W
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.;
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton
A;Title: Complete genome sequence of a multiple drug resistant Salmone
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AH0692
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-230 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD01913.1; PID:g16502755; GSPDB
C;Genetics:
A;Gene: ydgQ
C;Superfamily: conserved hypothetical protein HI1688

Query Match   2.5%;   Score 7;   DB 2;   Length 230;
Best Local Similarity   100.0%;   Pred. No. 1e+02;
Matches   7;   Conservative   0;   Mismatches   0;   Indels   0;   G

Qy   60 ALGLGLA 66
           |||||||
Db   38 ALGLGLA 44

RESULT 55
B64920
probable membrane protein ydgQ - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar
C;Accession: B64920
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland,
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: B64920
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-231 <BLAT>
A;Cross-references: GB:AE000258; GB:U00096; NID:g2367121; PIDN:AAC74704
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: ydgQ
A;Start codon: GTG
C;Superfamily: conserved hypothetical protein HI1688
C;Keywords: transmembrane protein
F;18-34/Domain: transmembrane #status predicted <TM01>
F;38-54/Domain: transmembrane #status predicted <TM02>
F;71-87/Domain: transmembrane #status predicted <TM03>
F;97-113/Domain: transmembrane #status predicted <TM04>

; S12614; MUID:91057136; PMID:2243790
4
RNA
<CAT>
wens, P.C.; McNeil, K.A.; Wallace, J.C.; Ballard, F.J.
, 681-687, 1989
ion, amino acid sequences and assay cross-reactivities of porcine insu
; A60738; MUID:90039035; PMID:2809477
3
rotein
'X', 81-91 <FRA>
ilin
nal sequence #status predicted <SIG>
nsulin-like growth factor II #status experimental <MAT>
arboxyl-terminal propeptide (E peptide) #status predicted <CTP>
5/Disulfide bonds: #status predicted

rity   2.5%;   Score 7;   DB 2;   Length 181;
onservative   100.0%;   Pred. No. 85;   Indels   0;   Gaps   0;
nservative   0;   Mismatches   0;   Indels   0;

AL 61
AL 18

uke protein [imported] - Arabidopsis thaliana
sis thaliana (mouse-ear cress)
#sequence_revision 16-Feb-2001 #text_change 16-Feb-2001

ropean Union Arabidopsis Genome Sequencing Consortium, The Cold Sprin
, 1999
and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A85001; MUID:20083488; PMID:10617198
ry
IA
STO:
GB:NC_001268; NID:g7269845; PIDN:CAB79704.1; GSPDB:GN00140

rity   2.5%;   Score 7;   DB 2;   Length 187;
              100.0%;   Pred. No. 87;
nservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

VS 78
VS 24

n C35D10.8 - Caenorhabditis elegans
bditis elegans
#sequence_revision 25-Apr-1997 #text_change 20-Sep-1999

ough, R.; Anderson, K.; Baynes, C.; Berks, M.; Bonfield, J.; Burton,
A.; Green, P.; Hawkins, T.; Hillier, L.; Jier, M.; Johnston, L.; Jon
M.; Parsons, J.; Percy, C.; Rifken, L.; Roopra, A.; Saunders, D.
BL Data Library, February 1995
n, R.; Smaldon, N.; Smith, A.; Sonnhammer, E.; Staden, R.; Sulston, J
, P.
C. elegans genome project: Contiguous nucleotide sequence of over two
S72566
IA
WIL:
EMBL:U21324; NID:g687879; PID:g687888
ce: strain Bristol N2

C86665
amino acid ABC trasporter permease protein [imported] - Lactococcus 1
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-De
C;Accession: C86665
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Wei
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium La
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: C86665
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-231 <STO>
A;Cross-references: GB:AE005176; PID:g12723189; PIDN:AAK04421.1; GSPD
A;Experimental source: strain IL1403
C;Genetics:
C;Gene: ydcC
C;Superfamily: ABC transporter permease protein

Query Match   2.5%; Score 7; DB 2; Length 231;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches   7; Conservative   0; Mismatches   0; Indels   0;

Qy   56 LVPLALG 62
           |||||||
Db   107 LVPLALG 113

RESULT 59
S60767
ribonuclease III - Coxiella burnetii
C;Species: Coxiella burnetii
C;Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 22-Ju
C;Accession: S60767
R;Zuber, M.; Hoover, T.A.; Powell, B.S.; Court, D.L.
Mol. Microbiol. 14, 291-300, 1994
A;Title: Analysis of the rnc locus of Coxiella burnetii.
A;Reference number: S60767; MUID:95131751; PMID:7830573
A;Accession: S60767
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-233 <ZUB>
A;Cross-references: EMBL:L27436; NID:g439870; PIDN:AAA69690.1; PID:g4
C;Superfamily: ribonuclease III; double-stranded RNA-binding repeat h
F;150-223/Domain: double-stranded RNA-binding repeat homology <DSR>

Query Match   2.5%; Score 7; DB 2; Length 233;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches   7; Conservative   0; Mismatches   0; Indels   0;

Qy   10 ARRLPLP 16
           |||||||
Db   164 ARRLPLP 170

RESULT 60
AB0273
probable membrane protein YPO2240 [imported] - Yersinia pestis (strai
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov
C;Accession: AB0273
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davi
il M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitel
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of pl
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AB0273
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-233 <KUR>
A;Cross-references: GB:AL590842; PIDN:CAC91046.1; PID:g15980240; GSPDI
C;Genetics:

---

transmembrane #status predicted <TM05>
transmembrane #status predicted <TM06>

larity 2.5%; Score 7; DB 2; Length 231;
       100.0%; Pred. No. 1e+02;
Conservative   0; Mismatches   0; Indels   0; Gaps   0;

LGLA 66
|||
LGLA 44

ein ECs2341 [imported] - Escherichia coli (strain O157:H7, substrain RI
ichia coli
01 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
21
kino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
unaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
, 2001
genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
21
r: A99629; MUID:21156231; PMID:11258796
nary
DNA
<HAY>
s: GB:BA000007; PIDN:BAB35764.1; PID:g13361808; GSPDB:GN00154
urce: strain O157:H7, substrain RIMD 0509952

nserved hypothetical protein HI1688

larity 2.5%; Score 7; DB 2; Length 231;
       100.0%; Pred. No. 1e+02;
Conservative   0; Mismatches   0; Indels   0; Gaps   0;

LGLA 66
|||
LGLA 44

ein ydgQ [imported] - Escherichia coli (strain O157:H7, substrain EDL93
ichia coli
01 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
70
nkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
ck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
83, 2001
sequence of enterohemorrhagic Escherichia coli O157:H7.
70
nary
DNA
<STO>
s: GB:AE005174; NID:g12515621; PIDN:AAG56621.1; GSPDB:GN00145; UWGP:226
urce: strain O157:H7, substrain EDL933

nserved hypothetical protein HI1688

larity 2.5%; Score 7; DB 2; Length 231;
       100.0%; Pred. No. 1e+02;
Conservative   0; Mismatches   0; Indels   0; Gaps   0;

LGLA 66
|||
LGLA 44

erved hypothetical protein HI1688

rity   2.5%; Score 7; DB 2; Length 233;
  100.0%; Pred. No. 1.1e+02;
nservative   0; Mismatches   0; Indels   0; Gaps   0;

LA 66

LA 44

synthase-like protein [imported] - Arabidopsis thaliana
sis thaliana (mouse-ear cress)
#sequence_revision 16-Feb-2001 #text_change 17-May-2002
ropean Union Arabidopsis Genome Sequencing Consortium, The Cold Sprin
1999
nd analysis of chromosome 4 of the plant Arabidopsis thaliana.
A85001; MUID:20083488; PMID:10617198
ry
A
STO>
GB:NC_001268; NID:g7267660; PIDN:CAB78088.1; GSPDB:GN00140

ransporting ATP synthase delta chain

rity   2.5%; Score 7; DB 2; Length 234;
  100.0%; Pred. No. 1.1e+02;
nservative   0; Mismatches   0; Indels   0; Gaps   0;

SS 237
||
SS 54

nslocating NADH dehydrogenase (ubiquinone) (EC 1.6.5.-) nqrD chain HI
lus influenzae
#sequence_revision 18-Aug-1995 #text_change 21-Jul-2000

; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
ott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
2, 1995
.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
me random sequencing and assembly of Haemophilus influenzae Rd.
A64000; MUID:95350630; PMID:7542800

cid sequence not shown; translation not shown
A
TIGR>
GB:U32841; GB:L42023; NID:g1574529; PIDN:AAC23334.1; PID:g1574540;
erved hypothetical protein HI1688
ductase

rity   2.5%; Score 7; DB 2; Length 235;
  100.0%; Pred. No. 1.1e+02;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

LA 66
||
LA 68

---

probable cytochrome b-561 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-
C;Accession: A86387
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; W
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Kha
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Mai
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.;
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: A86387
A;Accession: A86387
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-236 <STO>
A;Cross-references: GB:AE005172; NID:g11079498; PIDN:AAG29209.1; GSPDB:
C;Genetics:
A;Map position: 1

Query Match   2.5%; Score 7; DB 2; Length 236;
Best Local Similarity   100.0%; Pred. No. 1.1e+02;
Matches   7; Conservative   0; Mismatches   0; Indels   0; Ga

Qy   61 LGLGLAL 67
            |||||||
Db   194 LGLGLAL 200

RESULT 64
H70866
hypothetical protein Rv2473 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-
C;Accession: H70866
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harr
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlir
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squa
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrel
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: H70866
A;Status: preliminary; nucleic acid sequence not shown; translation not
A;Molecule type: DNA
A;Residues: 1-238 <COL>
A;Cross-references: GB:AL021246; GB:AL123456; NID:g3261507; PIDN:CAA16(
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: Rv2473

Query Match   2.5%; Score 7; DB 2; Length 238;
Best Local Similarity   100.0%; Pred. No. 1.1e+02;
Matches   7; Conservative   0; Mismatches   0; Indels   0; Ga

Qy   74 LAVVSLG 80
            |||||||
Db   90 LAVVSLG 96

RESULT 65
G83208
conserved hypothetical protein PA3494 [imported] - Pseudomonas aerugino
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 27-Nov-
C;Accession: G83208
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000

genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
.A82950; MUID:20437337; PMID:10984043
...ary
..NA
<STO>
.:GB:AE004770; GB:AE004091; NID:g9949633; PIDN:AAG06882.1; GSPDB:GN001
cce: strain PA01

served hypothetical protein HI1688

arity   2.5%;   Score 7; DB 2; Length 240;
.. 100.0%; Pred. No. 1.1e+02;
onservative 0; Mismatches 0; Indels 0; Gaps 0;

GLA 66
|||
GLA 46

ynthesis protein moeB - Haemophilus influenzae (strain Rd KW20)
ilus influenzae
5 #sequence_revision 18-Aug-1995 #text_change 29-Sep-1999
4
.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
cott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, S
.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
12, 1995
C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
ome random sequencing and assembly of Haemophilus influenzae Rd.
: A64000; MUID:95350630; PMID:7542800
4
acid sequence not shown; translation not shown
NA
<TIGR>
:: GB:U32823; GB:L42023; NID:g1574281; PIDN:AAC23099.1; PID:g1574288;

opterin biosynthesis
ybdopterin biosynthesis protein moeB
lenum; molybdopterin biosynthesis

arity   2.5%;   Score 7; DB 2; Length 243;
. 100.0%; Pred. No. 1.1e+02;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

NRQI 185
||||
NRQI 16

ein ZC196.8 - Caenorhabditis elegans
habditis elegans
99 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
42

EMBL Data Library, April 1997
e sequence of C. elegans cosmid ZC196.
r: Z20115
42
nary; translated from GB/EMBL/DDBJ
DNA
<MUR>
s: EMBL.U97007; PIDN:AAB52298.1; GSPDB:GN00023; CESP:ZC196.8
urce: strain Bristol N2; clone ZC196

6.8

A;Map position: 5
A;Introns: 41/1; 84/2; 144/3; 195/3

Query Match   2.5%;   Score 7; DB 2; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches   7; Conservative 0; Mismatches 0; Indels 0; G

Qy   234 AASSLGP 240
         |||||||
Db   225 AASSLGP 231

RESULT 68
A46066
lymphotoxin beta - human
C;Species: Homo sapiens (man)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul
C;Accession: A46066
R;Browning, J.L.; Ngam-ek, A.; Lawton, P.; DeMarinis, J.; Tizard, R.;
Cell 72, 847-856, 1993
A;Title: Lymphotoxin beta, a novel member of the TNF family that forms
A;Reference number: A46066; MUID:93208881; PMID:7916655
A;Accession: A46066
A;Status: preliminary
A;Molecule type: DNA; protein
A;Residues: 1-244 <BRO>
A;Cross-references: GB:L11015; NID:g292276; PIDN:AAA36191.1; PID:g2922
A;Note: sequence extracted from NCBI backbone (NCBIN:128066, NCBIP:128
C;Keywords: transmembrane protein

Query Match   2.5%;   Score 7; DB 2; Length 244;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches   7; Conservative 0; Mismatches 0; Indels 0; C

Qy   194 GLYYLYC 200
         |||||||
Db   131 GLYYLYC 137

RESULT 69
B32352
molybdopterin biosynthesis protein moeB - Escherichia coli (strain K-
N;Alternate names: molybdopterin-converting factor chlN
C;Species: Escherichia coli
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 01-Ma
C;Accession: B32352; B64820
R;Nohno, T.; Kasai, Y.; Saito, T.
J. Bacteriol. 170, 4097-4102, 1988
A;Title: Cloning and sequencing of the Escherichia coli chlEN operon
A;Reference number: A32352; MUID:88314906; PMID:3045084
A;Accession: B32352
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-249 <NOH>
A;Cross-references: GB:M21151; NID:g145538; PIDN:AAA23580.1; PID:g145
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: B64820
A;Status: preliminary; nucleic acid sequence not shown; translation n
A;Molecule type: DNA
A;Residues: 1-249 <BLAT>
A;Cross-references: GB:AE000185; GB:U00096; NID:g1787047; PIDN:AAC739
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: moeB; chlN
A;Map position: 18 min
C;Function:
C;Pathway: molybdopterin biosynthesis
C;Superfamily: molybdopterin biosynthesis protein moeB
C;Keywords: molybdenum; molybdopterin biosynthesis

F

C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; W
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.;
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton
A;Title: Complete genome sequence of a multiple drug resistant Salmone
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AI0602
A;Accession: AI0602
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-249 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD05291.1; PID:gi16502055; GSPDE
C;Genetics:
A;Gene: STY0884
C;Superfamily: molybdopterin biosynthesis protein moeB

Query Match　　　　2.5%;　Score 7;　DB 2;　Length 249;
Best Local Similarity　　100.0%;　Pred. No. 1.1e+02;
Matches　　7;　Conservative　　0;　Mismatches　　0;　Indels　　0;　C

Qy　　179 LRYNRQI 185
　　　　　|||||||
Db　　　10 LRYNRQI 16

RESULT 73
AI3154
IS21 family transposase istB [imported] - Agrobacterium tumefaciens (s
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov
C;Accession: AI3154
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy,
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tum
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AI3154
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-252 <KUR>
A;Cross-references: GB:AE008689; PIDN:AAL45655.1; PID:gi17743380; GSPD
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: istB
A;Map position: linear chromosome
C;Superfamily: DNA replication protein dnaC

Query Match　　　　2.5%;　Score 7;　DB 2;　Length 252;
Best Local Similarity　　100.0%;　Pred. No. 1.1e+02;
Matches　　7;　Conservative　　0;　Mismatches　　0;　Indels　　0;

Qy　　60 ALGLGLA 66
　　　　|||||||
Db　　120 ALGLGLA 126

RESULT 74
H98132
istB protein (AJ238712) [imported] - Agrobacterium tumefaciens (strai
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-No
C;Accession: H98132
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Q
R.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappi
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Age

---

arity　　　2.5%;　Score 7;　DB 2;　Length 249;
onservative　　0;　Mismatches　　0;　Indels　　0;　Gaps　　0;

RQI 185
|||
RQI 16

ynthesis MoeB protein [imported] - Escherichia coli (strain O157:H7, s
chia coli
1 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
ino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
naga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
2001
genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
: A99629; MUID:21156231; PMID:11258796
ary
NA
<HAY>
: GB:BA000007; PIDN:BAB34327.1; PID:gi13360363; GSPDB:GN00154
rce: strain O157:H7, substrain RIMD 0509952

ybdopterin biosynthesis protein moeB

arity　　　2.5%;　Score 7;　DB 2;　Length 249;
onservative　　0;　Mismatches　　0;　Indels　　0;　Gaps　　0;

RQI 185
|||
RQI 16

ynthesis [imported] - Escherichia coli (strain O157:H7, substrain EDL9
ichia coli
01 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
92
ankett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
ck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
33, 2001
equence of enterohemorrhagic Escherichia coli O157:H7.
r: A85480; MUID:21074935; PMID:11206551
92
nary
DNA
<STO>
B: GB:AE005174; NID:gi12513829; PIDN:AAG55199.1; GSPDB:GN00145; UWGP:210
urce: strain O157:H7, substrain EDL933

lybdopterin biosynthesis protein moeB

larity　　　2.5%;　Score 7;　DB 2;　Length 249;
Conservative　　0;　Mismatches　　0;　Indels　　0;　Gaps　　0;

NRQI 185
||||
NRQI 16

osynthesis MoeB protein [imported] - Salmonella enterica subsp. enterica

c: A97359; MUID:21608551; PMID:11743194
32
nary
DNA
<KUR>
s: GB:AE007870; PIDN:AAK88586.1; PID:g15158297; GSPDB:GN00170

inear chromosome
A replication protein dnaC

larity   2.5%;   Score 7;   DB 2;   Length 252;
100.0%;   Pred. No. 1.1e+02;
Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

LGLA 66
||||
LGLA 126

synthesis MoeB protein VCA0618 [imported] - Vibrio cholerae (strain N16
cholerae
00 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
39
.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
olaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, E
s, J.J.; Venter, J.C.; Fraser, C.M.
B3, 2000
ence of both chromosomes of the cholera pathogen Vibrio cholerae.
r: AB2035; MUID:20406833; PMID:10952301
39
nary
DNA
<HEI>
s: GB:AE004391; GB:AE003853; NID:g9658015; PIDN:AAF96519.1; GSPDB:GN001
urce: serogroup O1; strain N16961; biotype El Tor

lybdopterin biosynthesis protein moeB

larity   2.5%;   Score 7;   DB 2;   Length 254;
100.0%;   Pred. No. 1.1e+02;
Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

NRQI 185
||||
NRQI 18

April 7, 2004, 17:59:25
s

```
        GenCore version 5.1.6
opyright (c) 1993 - 2004 Compugen Ltd.

.n search, using sw model

:il  7, 2004, 17:48:13 ; Search time 17 Seconds
                         (without alignments)
                 869.878 Million cell updates/sec

-09-245-198A-4
.SLLDFEISARRLPLPRSLG.........PWAHLKAAPFLTYFGLFQVH 284

:GO
.op 60.0 , Gapext 60.0
.681 seqs, 52070155 residues

.s satisfying chosen parameters:      141681

.gth: 0
.gth: 2000000000

.sting first 100 summaries

wissProt_42:*

 the number of results predicted by chance to have a
: than or equal to the score of the result being printed,
:d by analysis of the total score distribution.
```

SUMMARIES

ary	ch	Length	DB	ID	Description
7.7		249	1	TN12_HUMAN	O43508 homo sapien
1.3		225	1	TN12_MOUSE	O54907 mus musculu
3.5		111	1	YFBW_ECOLI	Q47377 escherichia
1.2		733	1	PSAB_ODOSI	P49480 odontella s
3.2		734	1	PSAB_CYACA	Q9tlq6 cyanidium c
2.8		179	1	ADHS_GLUOX	O05544 gluconobact
2.8		220	1	Y304_BRUME	Q8yd73 brucella me
2.8		220	1	Y3J1_BRUSU	Q8fv59 brucella su
2.8		317	1	MSHR_PANTR	Q9tuk4 pan troglod
2.8		379	1	FDHB_METJA	Q60316 methanococc
2.8		422	1	ZP3_MESAU	P23491 mesocricetu
2.8		576	1	CYDC_HAEIN	P45081 haemophilus
2.8		586	1	GGT5_HUMAN	P36269 homo sapien
2.5		24	1	HSP3_OCTVU	P83215 octopus vul
2.5		30	1	HSP5_OCTVU	P83217 octopus vul
2.5		49	1	HSP1_SAGIM	P24714 saguinus im
2.5		57	1	HSP1_DIDMA	P35305 didelphis m
2.5		115	1	A62F_DROME	O46202 drosophila
2.5		118	1	YI51_ECOLI	P03838 escherichia
2.5		131	1	IL13_MOUSE	P20109 mus musculu
2.5		131	1	IL13_RAT	P42203 rattus norv
2.5		147	1	YK01_PYRHO	O57781 pyrococcus
2.5		150	1	TNFC_PIG	Q9t6v8 sus scrofa
2.5		157	1	RA05_ORYSA	Q01881 oryza sativ
2.5		161	1	HLPA_ECOLI	P11457 escherichia
2.5		170	1	YLM4_CAEEL	P34378 caenorhabdi
2.5		180	1	LACB_BUBBU	P02755 bubalus bub
2.5		180	1	LACB_CAPHI	P02756 capra hircu
2.5		180	1	LACB_SHEEP	P02757 ovis aries
2.5		181	1	APT_SHEON	Q8efg1 shewanella
2.5		217	1	DEFI_BIFLO	Q8g534 bifidobacte
2.5		230	1	RNFE_SALTY	Q8xex9 salmonella
2.5		231	1	RNFE_ECO57	P58344 escherichia

#	score		Length	DB	ID	accession	org
34	2.5	7	231	1	RNFE_ECOLI	P77179	esch
35	2.5	7	233	1	RNC_COXBU	P51837	coxi
36	2.5	7	233	1	RNFE_YERPE	Q8zed4	yers
37	2.5	7	235	1	RNFE_HAEIN	Q57020	haem
38	2.5	7	239	1	TN14_MOUSE	Q9qyh9	mus
39	2.5	7	240	1	RNFE_PSEAE	Q9hyb5	pseu
40	2.5	7	243	1	MOEB_HAEIN	P45211	haem
41	2.5	7	244	1	TNFC_HUMAN	Q06643	homo
42	2.5	7	244	1	TNFC_PANTR	Q862z7	pan
43	2.5	7	249	1	LFTR_XANCP	Q8p996	xant
44	2.5	7	249	1	MOEB_ECOLI	P12282	esch
45	2.5	7	249	1	MOEB_SALTY	Q56067	salm
46	2.5	7	257	1	KDTX_SERMA	Q54435	serr
47	2.5	7	277	1	CN09_HUMAN	Q86t03	homo
48	2.5	7	310	1	TNFC_MARMO	Q9jm10	marm
49	2.5	7	316	1	ISPH_XANCP	Q8pbg4	xant
50	2.5	7	316	1	ISPH_XYLFA	Q9pas9	xyle
51	2.5	7	324	1	ODPB_BACSU	P21882	baci
52	2.5	7	329	1	SRA6_CAEEL	Q09208	caen
53	2.5	7	335	1	IAG2_RAT	O35777	ratt
54	2.5	7	344	1	LEU3_THEAQ	P24098	ther
55	2.5	7	352	1	LEU3_DEIRA	Q9rth9	dein
56	2.5	7	357	1	G6PT_MOUSE	P35576	mus
57	2.5	7	357	1	G6PT_RAT	P43428	ratt
58	2.5	7	358	1	PON1_RABIT	P27170	oryc
59	2.5	7	365	1	NQO8_THETH	Q60019	ther
60	2.5	7	394	1	BENE_ACICA	P07775	acin
61	2.5	7	396	1	DHH1_XENLA	Q91610	xeno
62	2.5	7	398	1	DHH2_XENLA	Q91611	xeno
63	2.5	7	402	1	SELP_BOVIN	P49907	bos
64	2.5	7	412	1	PGKP_ALCEU	P50320	alca
65	2.5	7	413	1	PGKC_ALCEU	P50319	alca
66	2.5	7	416	1	NH59_CAEEL	Q9txj1	caen
67	2.5	7	418	1	CPI6_RAT	P09006	ratt
68	2.5	7	419	1	ENO_PYRAE	Q8zye7	pyro
69	2.5	7	419	1	PGK_RALSO	Q8ylw6	rals
70	2.5	7	423	1	YJ54_YEAST	P47130	sacc
71	2.5	7	424	1	ZP3_MOUSE	P10761	mus
72	2.5	7	428	1	SYH_CHLMU	Q9pjj9	chla
73	2.5	7	455	1	PHR_STRGR	P12768	stre
74	2.5	7	461	1	PUCC_RHOCA	P23462	rhod
75	2.5	7	461	1	Y608_HAEIN	Q57486	haem
76	2.5	7	464	1	SOX8_MOUSE	Q04886	mus
77	2.5	7	467	1	D4DR_HUMAN	P21917	homo
78	2.5	7	483	1	GLME_CLOCO	P80077	clos
79	2.5	7	485	1	GLME_CLOTT	Q05509	clos
80	2.5	7	505	1	PDI_HUMIN	P55059	humi
81	2.5	7	512	1	LNT_ECOLI	P23930	esch
82	2.5	7	512	1	LNT_SALTY	O87576	salm
83	2.5	7	518	1	ASB3_HUMAN	Q9y575	homo
84	2.5	7	525	1	ASB3_MOUSE	Q9wv72	mus
85	2.5	7	547	1	RM56_HUMAN	P83111	homo
86	2.5	7	551	1	RM56_MOUSE	Q9ep89	mus
87	2.5	7	600	1	S133_MOUSE	Q91y63	mus
88	2.5	7	600	1	S133_RAT	Q9z0z5	ratt
89	2.5	7	653	1	MUTL_VIBCH	Q9kv13	vibr
90	2.5	7	720	1	FPTA_PSEAE	P42512	pseu
91	2.5	7	758	1	CSTA_MYCTU	P95095	myco
92	2.5	7	760	1	PO21_XENLA	P16143	xeno
93	2.5	7	810	1	NEL1_HUMAN	Q92832	homo
94	2.5	7	810	1	NEL1_RAT	Q62919	ratt
95	2.5	7	854	1	KDPD_RATRA	O34971	rath
96	2.5	7	953	1	B3A4_RAT	Q8k4v2	ratt
97	2.5	7	955	1	B3A4_RABIT	Q9gky1	oryc
98	2.5	7	983	1	B3A4_HUMAN	Q96q91	homo
99	2.5	7	1072	1	ITA6_CHICK	P26007	gall
100	2.5	7	1076	1	IF3A_CAEEL	P34339	caen

ALIGNMENTS

RESULT 1

```
          STANDARD;      PRT;     249 AA.
7;  Rel. 41, Created)
    Rel. 41, Last sequence update)
    Rel. 42, Last annotation update)
s factor ligand superfamily member 12 (TNF-related weak
poptosis) (TWEAK) (APO3 ligand).
O3L OR DR3LG.
(Human).
tazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
heria; Primates; Catarrhini; Hominidae; Homo.
606;
! N.A., AND N-TERMINUS OF SOLUBLE FORM.
liver, and Tonsil;
415; PubMed=9405449;
ne Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H.,
arcia I., Browning J.L.;
secreted ligand in the tumor necrosis factor family that
s apoptosis.";
. 272:32401-32410(1997).
! N.A.
kidney;
355; PubMed=9560343;
, Sheridan J.P., Pitti R.M., Brush J., Goddard A.,
on of a ligand for the death-domain-containing receptor
3:525-528(1998).
! N.A.
257; PubMed=12477932;
L., Feingold E.A., Grouse L.H., Derge J.G.,
, Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
quellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
cEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
ton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
adan E., Young A.C., Shevchenko Y., Bouffard G.G.,
, Touchman J.W., Green E.D., Dickson M.C.,
, Grimwood J., Schmutz J., Myers R.M.,
.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
Schein J.E., Jones S.J.M., Marra M.A.;
nd initial analysis of more than 15,000 full-length
use cDNA sequences.";
cad. Sci. U.S.A. 99:16899-16903(2002).
061; PubMed=10085077;
ang Y.C., Lund J.K., Chen Y.-W., Leal J.A., Wiley S.R.;
s angiogenesis and proliferation of endothelial cells.";
. 274:8455-8459(1999).
Binds to FN14 and possibly also to TNRFSF12/APO3. Weak
of apoptosis in some cell types. Mediates NF-KappaB
n. May promote angiogenesis and the proliferation of
al cells.
Homotrimer (Potential).
AR LOCATION: Type II membrane protein and secreted.
ECIFICITY: Highly expressed in adult heart, pancreas,
muscle, brain, colon, small intestine, lung, ovary,
spleen, lymph node, appendix and peripheral blood
es. Low expression in kidney, testis, liver, placenta,
d bone marrow. Also detected in fetal kidney, liver,
brain.
```

```
CC   -!- PTM: The soluble form derives from the membrane form
CC       by proteolytic processing.
CC   -!- SIMILARITY: Belongs to the tumor necrosis factor family.
CC   -!- CAUTION: Ref.3 sequence differs from that shown due to a
CC       frameshift in position 125.
CC   --------------------------------------------------------------
```

```
DR   EMBL; AF030099; AAC51923.1; -.
DR   EMBL; AF055872; AAC39724.1; -.
DR   EMBL; BC019047; AAH19047.1; ALT_FRAME.
DR   Genew; HGNC:11927; TNFSF12.
DR   MIM; 602695; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0005102; F:receptor binding; TAS.
DR   GO; GO:0006917; P:induction of apoptosis; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR006052; TNF family.
DR   InterPro; IPR008983; TNF_like.
DR   Pfam; PF00229; TNF; 1.
DR   SMART; SM00207; TNF; 1.
DR   PROSITE; PS00251; TNF_1; FALSE_NEG.
DR   PROSITE; PS50049; TNF_2; 1.
KW   Cytokine; Angiogenesis; Apoptosis; Transmembrane; Glycoprotein;
KW   Signal-anchor.
FT   CHAIN         1    249   TUMOR NECROSIS FACTOR LIGAND SUPERF/
FT                          MEMBER 12, MEMBRANE FORM.
FT   CHAIN        94    249   TUMOR NECROSIS FACTOR LIGAND SUPERF/
FT                          MEMBER 12, SECRETED FORM.
FT   DOMAIN        1     21   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     22     42   SIGNAL-ANCHOR (TYPE-II MEMBRANE PRO
FT                          (POTENTIAL).
FT   DOMAIN       43    249   EXTRACELLULAR (POTENTIAL).
FT   SITE         93     94   CLEAVAGE.
FT   CARBOHYD    139    139   N-LINKED (GLCNAC...).
SQ   SEQUENCE   249 AA;  27216 MW;  E6608433E1C28EBA CRC64;
```

```
Query Match             87.7%;  Score 249;  DB 1;  Length 249;
Best Local Similarity   100.0%;  Pred. No. 2.9e-232;
Matches  249;  Conservative    0;  Mismatches    0;  Indels    0;  (

Qy    36 MAARRSQRRRGRRGBEPGTALLVPLALGLGLALACLGLLLAVVSLGSRASLSAQEP/
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MAARRSQRRRGRRGBEPGTALLVPLALGLGLALACLGLLLAVVSLGSRASLSAQEP/

Qy    96 VAEEDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPI
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 VAEEDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPI

Qy   156 GAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAV\
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAV\

Qy   216 LLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLK/
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 LLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLK/

Qy   276 TYFGLFQVH 284
         |||||||||
Db   241 TYFGLFQVH 249

RESULT 2
TN12_MOUSE
ID   TN12_MOUSE     STANDARD;      PRT;   225 AA.
AC   O54907; Q9CTP2;
DT   28-FEB-2003 (Rel. 41, Created)
```

FT   NON_TER   1   1
FT   CHAIN   <1   225   TUMOR NECROSIS FACTOR LIGAND SUPERFA
FT   MEMBER 12, MEMBRANE FORM.
FT   CHAIN   70   225   TUMOR NECROSIS FACTOR LIGAND SUPERFA
FT   MEMBER 12, SECRETED FORM (BY SIMILAR
FT   TRANSMEM   <1   21   SIGNAL-ANCHOR (TYPE-II MEMBRANE PROT
FT   (POTENTIAL).
FT   DOMAIN   22   225   EXTRACELLULAR (POTENTIAL).
FT   SITE   69   70   CLEAVAGE (BY SIMILARITY).
FT   DISULFID   167   186   POTENTIAL.
FT   CARBOHYD   115   115   N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   225 AA; 24781 MW; 90C412CC0480659B CRC64;

Query Match   11.3%;
Best Local Similarity   100.0%;   Score 32; DB 1; Length 225;
Matches   32;   Conservative   0;   Mismatches   0;   Indels   0;   G

Qy   139 RRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEE 170
         ||||||||||||||||||||||||||||||||
Db   80 RRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEE 111

RESULT 3
YFBW_ECOLI
ID   YFBW_ECOLI   STANDARD;   PRT;   111 AA.
AC   Q47377;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical protein yfbW.
GN   YFBW OR B2257.1.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE OF 81-111 FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=96186953; PubMed=8626063;
RA   Sharma V., Hudspeth M.E., Meganathan R.;
RT   "Menaquinone (vitamin K2) biosynthesis: localization and
RT   characterization of the menE gene from Escherichia coli.";
RL   Gene 168:43-48(1996).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   ----------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------
DR   EMBL; AE000315; -; NOT_ANNOTATED_CDS.
DR   EMBL; L35031; AAB04895.1; -.
DR   EcoGene; EG14344; yfbW.
DR   InterPro; IPR000620; DUF6.
DR   Pfam; PF00892; DUF6; 1.
KW   Hypothetical protein; Transmembrane; Complete proteome.
FT   TRANSMEM   36   56   POTENTIAL.
FT   TRANSMEM   62   82   POTENTIAL.
FT   TRANSMEM   88   108   POTENTIAL.
SQ   SEQUENCE   111 AA; 12192 MW; 7CFA06D75DA33D69 CRC64;

(Rel. 41, Last sequence update)
(Rel. 41, Last annotation update)
...s factor ligand superfamily member 12 (TNF-related weak
...ptosis) (TWEAK) (Fragment).
(Mouse).
...azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
...teria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
...090;
N.A.
...eal macrophage;
...115; PubMed=9405449;
...rcia I., Bourdon P.R., Xu H., Hsu Y.-M., Scott H.,
...Y., Browning J.L.;
...secreted ligand in the tumor necrosis factor family that
...apoptosis.";
...272:32401-32410(1997).

...-225 FROM N.A.
...J; TISSUE=Retina;
...60; PubMed=11217851;
...nagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
...ira A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
...wa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
...suda H.A., Ashburner M., Batalov S., Casavant T.,
..., Gaasterland T., Gissi C., King B., Kochiwa H.,
...s S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
...Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
...io T., Furuno M., Aono H., Baldarelli R., Barsh G.,
...elli D., Bojunga N., Carninci P., de Bonaldo M.F.,
..., Bult C., Fletcher C., Fujita M., Gariboldi M.,
...Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
...chionni L., Mashima J., Mazzarelli J., Mombaerts P.,
...ng B., Ringwald M., Rodriguez I., Sakamoto N.,
...o K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
...vo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
...A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
...;
...notation of a full-length mouse cDNA collection.";
...-690(2001).
...Binds to FN14 and possibly also to TNRFSF12/APO3. Weak
...apoptosis in some cell types. Promotes angiogenesis and
...eration of endothelial cells. Mediates NF-KappaB
...(By similarity).
...omotrimer (Potential).
...R LOCATION: Type II membrane protein and secreted (By
...).
...ECIFICITY: Widely expressed.
...soluble form is produced from the membrane form by
...c processing (By similarity).
...: Belongs to the tumor necrosis factor family.
...----------------
...OT entry is copyright. It is produced through a collaboration
...Swiss Institute of Bioinformatics and the EMBL outstation -
...Bioinformatics Institute. There are no restrictions on its
...profit institutions as long as its content is in no way
...his statement is not removed. Usage by and for commercial
...res a license agreement (See http://www.isb-sib.ch/announce/
...il to license@isb-sib.ch).
...----------------
...D; AAC53517.1; -.
...3; BAB32249.1; -.
...259; Tnfsf12.
...D06052; TNF_family.
...008983; TNF_like.
...; TNF; 1.
...7; TNF; 1.
...251; TNF_1; FALSE_NEG.
...049; TNF_2; 1.
...logenesis; Apoptosis; Transmembrane; Glycoprotein;

```
larity     3.5%;   Score 10;  DB 1;  Length 111;
onservative 100.0%; Pred. No. 0.038;  Indels  0;  Mismatches  0;  Gaps  0;

ALACLGL 72
|||||||
ALACLGL 49
```

```
        STANDARD;     PRT;   733 AA.
(Rel. 33, Created)
(Rel. 33, Last sequence update)
(Rel. 42, Last annotation update)
I P700 chlorophyll A apoprotein A2 (PsaB) (PSI-B).
nensis (Marine centric diatom).
cramenopiles; Bacillariophyta; Coscinodiscophyceae;
ycidae; Eupodiscales; Eupodiscaceae; Odontella.
339;
M N.A.
., Stoebe B., Schaffran I., Kroth-Pancic P., Freier U.;
last genome of a chlorophyll a+c-containing alga,
nensis.";
iol. Rep. 13:336-342(1995).
: PsaA and psaB bind P700, the primary electron donor of
tem I (PSI), as well as the electron acceptors A0, A1 and
copically characterized acceptors A0, A1, FX, FA and FB in
idized P700 is reduced on the lumenal side of the
ng photonic excitation into a charge separation, which
s an electron from the donor P700 chlorophyll pair to the
is a plastocyanin/cytochrome c6-ferredoxin oxidoreductase,
d membrane by plastocyanin or cytochrome c6.
; P700 is a chlorophyll A/chlorophyll A' dimer, A0 is one
chlorophyll A, A1 is one or both phylloquinones and FX is
4Fe-4S iron-sulfur center (By similarity).
The psaA/B heterodimer binds the P700 chlorophyll special
subsequent electron acceptors. PSI consists of a core
complex that captures photons, and an electron transfer
at converts photonic excitation into a charge separation.
ryotic PSI reaction center is composed of at least 11
(By similarity).
LAR LOCATION: Integral membrane protein. Chloroplast
d membrane (By similarity).
TY: Belongs to the psaA/psaB family.
```

```
------------------------------------------------
; CAA91749.1; -.
S78376.
; 1JB0.
482; -; 1.
R006244; PsaB.
R001280; PSI_PsaA/B.
3; psaA_psaB; 1.
257; PHOTSYSPSAAB.
GR01336; psaB; 1.
0419; PHOTOSYSTEM I PSAAB; 1.
lectron transport; Photosynthesis; Thylakoid;
I; Chlorophyll; Metal-binding; Iron; Magnesium;
4Fe-4S; Transmembrane; Chloroplast.
46      69      I (POTENTIAL).
```

```
FT   TRANSMEM    134    157    II (POTENTIAL).
FT   TRANSMEM    174    198    III (POTENTIAL).
FT   TRANSMEM    272    290    IV (POTENTIAL).
FT   TRANSMEM    329    352    V (POTENTIAL).
FT   TRANSMEM    368    394    VI (POTENTIAL).
FT   TRANSMEM    416    438    VII (POTENTIAL).
FT   TRANSMEM    516    534    VIII (POTENTIAL).
FT   TRANSMEM    574    595    IX (POTENTIAL).
FT   TRANSMEM    642    664    X (POTENTIAL).
FT   TRANSMEM    706    726    XI (POTENTIAL).
FT   METAL       558    558    IRON-SULFUR (4FE-4S) (SHARED WITH D
FT                             PARTNER) (BY SIMILARITY).
FT   METAL       567    567    IRON-SULFUR (4FE-4S) (SHARED WITH D
FT                             PARTNER) (BY SIMILARITY).
FT   METAL       653    653    MAGNESIUM (CHLOROPHYLL-A B1 AXIAL L
FT                             P700 SPECIAL PAIR) (BY SIMILARITY).
FT   METAL       661    661    MAGNESIUM (CHLOROPHYLL-A B3 AXIAL L
FT                             (BY SIMILARITY).
FT   BINDING     669    669    CHLOROPHYLL-A B3 (BY SIMILARITY).
FT   BINDING     670    670    PHYLLOQUINONE B (BY SIMILARITY).
SQ   SEQUENCE   733 AA;  82103 MW;  13439AF1E441BBF7 CRC64;

Query Match             3.2%;  Score 9;  DB 1;  Length 733;
Best Local Similarity 100.0%;  Pred. No. 1.8;
Matches   9;  Conservative   0;  Mismatches   0;  Indels   0;

Qy      63 LGLAIACLG 71
           |||||||||
Db     333 LGLAIACLG 341
```

```
RESULT 5
PSAB_CYACA
ID   PSAB_CYACA      STANDARD;      PRT;     734 AA.
AC   Q9TLQ6;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Photosystem I P700 chlorophyll A apoprotein A2 (PsaB) (PSI-B).
GN   PSAB.
OS   Cyanidium caldarium.
OG   Chloroplast.
OC   Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridi
OC   Cyanidium.
OX   NCBI_TaxID=2771;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RK-1;
RX   MEDLINE=20496959; PubMed=11040290;
RA   Gloeckner G., Rosenthal A., Valentin K.-U.;
RT   "The structure and gene repertoire of an ancient red algal plast
RT   genome.";
RL   J. Mol. Evol. 51:382-390(2000).
CC   -!- FUNCTION: PsaA and psaB bind P700, the primary electron dono
CC       photosystem I (PSI), as well as the electron acceptors A0, A
CC       FX. PSI is a plastocyanin/cytochrome c6-ferredoxin oxidoredu
CC       converting photonic excitation into a charge separation, whi
CC       transfers an electron from the donor P700 chlorophyll pair t
CC       spectroscopically characterized acceptors A0, A1, FX, FA and
CC       turn. Oxidized P700 is reduced on the lumenal side of the
CC       thylakoid membrane by plastocyanin or cytochrome c6.
CC   -!- COFACTOR: P700 is a chlorophyll A/chlorophyll A' dimer, A0 i
CC       or more chlorophyll A, A1 is one or both phylloquinones and
CC       a shared 4Fe-4S iron-sulfur center (By similarity).
CC   -!- SUBUNIT: The psaA/B heterodimer binds the P700 chlorophyll s
CC       pair and subsequent electron acceptors. PSI consists of a co
CC       antenna complex that captures photons, and an electron trans
CC       chain that converts photonic excitation into a charge separa
CC       The eukaryotic PSI reaction center is composed of at least 1
CC       subunits (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast
CC       thylakoid membrane (By similarity).
CC   -!- SIMILARITY: Belongs to the psaA/psaB family.
```

RT   expression in Acetobacter pasteurianus.";
RL   Appl. Environ. Microbiol. 63:1131-1138(1997).
CC   -!- FUNCTION: NOT ESSENTIAL FOR ALCOHOL DEHYDROGENASE ACTIVITY.
CC   -!- SUBUNIT: HETEROTRIMER (DEHYDROGENASE, CYTOCHROME AND PROTEIN
CC   ADHS), THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX.
CC   -!- SUBCELLULAR LOCATION: Membrane-bound, facing the periplasmic
CC   (Potential).
CC   --
CC   --
DR   EMBL; D86440; BAA19756.1; -.
KW   Membrane; Periplasmic; Signal; Pyrrolidone carboxylic acid.
FT   SIGNAL   1   24   POTENTIAL.
FT   CHAIN   25   179   ALCOHOL DEHYDROGENASE 15 kDa SUBUNIT
FT   MOD_RES   25   25   PYRROLIDONE CARBOXYLIC ACID.
SQ   SEQUENCE   179 AA; 19943 MW; F6AF243656B3CC66 CRC64;

Query Match   2.8%;   Score 8; DB 1; Length 179;
Best Local Similarity   100.0%;   Pred. No. 4.9;
Matches   8; Conservative   0; Mismatches   0; Indels   0; G

Qy   59 LALGLGLA 66
            ||||||||
Db   11 LALGLGLA 18

RESULT 7
Y304_BRUME
ID   Y304_BRUME   STANDARD;   PRT;   220 AA.
AC   Q8YD73;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical protein BMEII0304.
GN   BMEII0304.
OS   Brucella melitensis.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Brucellaceae; Brucella.
OX   NCBI_TaxID=29459;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=16M / ATCC 23456 / Biotype 1;
RX   MEDLINE=20020109; PubMed=11756688;
RA   DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mujer C., Lo
RA   Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,
RA   Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsma
RA   Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA   Haselkorn R., Kyrpides N., Overbeek R.;
RT   "The genome sequence of the facultative intracellular pathogen
RT   Brucella melitensis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- SIMILARITY: Belongs to the UPF0191 family.
CC   --
CC   --
DR   EMBL; AE009669; AAL53546.1; -.
DR   PIR; AG3547; AG3547.
DR   HAMAP; MF_01207; -; 1.
DR   InterPro; IPR007916; UPF0191.
DR   Pfam; PF05252; UPF0191; 1.

...; AAF12881.1; -.
... 1UB0.
...2; -; 1.
...06244; PsaB.
...01280; PSI_PsaA/B.
...   psaA_psaB; 1.
...7; PHOTSYPSPSAAB.
...01336; PsaB; 1.
...19; PHOTOSYSTEM I PSAAB; 1.
...ctron transport; Photosynthesis; Thylakoid;
... Chlorophyll; Metal-binding; Iron; Magnesium;
...Fe-4S; Transmembrane; Chloroplast.
...6   69   I (POTENTIAL).
...5   158   II (POTENTIAL).
...5   199   III (POTENTIAL).
...3   291   IV (POTENTIAL).
...0   353   V (POTENTIAL).
...9   395   VI (POTENTIAL).
...7   439   VII (POTENTIAL).
...5   535   VIII (POTENTIAL).
...5   596   IX (POTENTIAL).
...3   665   X (POTENTIAL).
...7   727   XI (POTENTIAL).
...9   559   IRON-SULFUR (4FE-4S) (SHARED WITH DIMERIC
...   PARTNER) (BY SIMILARITY).
...8   568   IRON-SULFUR (4FE-4S) (SHARED WITH DIMERIC
...   PARTNER) (BY SIMILARITY).
...4   654   MAGNESIUM (CHLOROPHYLL-A B1 AXIAL LIGAND;
...   P700 SPECIAL PAIR) (BY SIMILARITY).
...2   662   MAGNESIUM (CHLOROPHYLL-A B3 AXIAL LIGAND)
...   (BY SIMILARITY).
...0   670   CHLOROPHYLL-A B3 (BY SIMILARITY).
...1   671   PHYLLOQUINONE B (BY SIMILARITY).
... AA; 82359 MW; 4496AA2AE59CA9B9 CRC64;

...   3.2%;   Score 9; DB 1; Length 734;
... 100.0%;   Pred. No. 1.8;
...nservative   0; Mismatches   0; Indels   0; Gaps   0;

...   ACLG 71
... ||||
...   ACLG 342

STANDARD;   PRT;   179 AA.

...el. 35, Created)
...el. 35, Last sequence update)
...el. 41, Last annotation update)
...ogenase 15 kDa subunit precursor (G3-ADH subunit III).

...oxydans (Gluconobacter suboxydans).
...eobacteria; Alphaproteobacteria; Rhodospirillales;
...ae; Gluconobacter.
...;

...N.A., AND SEQUENCE OF 26-40.
...528;
...225; PubMed=9055427;
...inouchi S.;
...tion of the genes encoding the three-component membrane-
...dehydrogenase from Gluconobacter suboxydans and their

Qy   54 ALLVPLAL 61
              ||||||||
Db   132 ALLVPLAL 139

RESULT 9
MSHR_PANTR
ID   MSHR_PANTR   STANDARD;   PRT;   317 AA.
AC   Q9TUK4; Q864L1;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Melanocyte stimulating hormone receptor (MSH-R) (Melanotropin
DE   receptor) (Melanocortin-1 receptor) (MC1-R).
GN   MC1R.
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rees J.L., Harding R.M., Healy E., Jackson I.J., Ray A.J., Ellis
RA   Flanagan N., Todd C., Dixon C., Matthews J.N., Sajantila A.,
RA   Birch-Machin M.A.;
RT   "Chimpanzee melanocortin 1 sequence.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Isolate 3;
RX   MEDLINE=22572539; PubMed=12687585;
RA   Mundy N.I., Kelly J.;
RT   "Evolution of a pigmentation gene, the melanocortin-1 receptor,
RT   primates.";
RL   Am. J. Phys. Anthropol. 121:67-80(2003).
CC   -!- FUNCTION: Receptor for MSH (alpha, beta and gamma) and ACTH.
CC   activity of this receptor is mediated by G proteins which act
CC   adenylate cyclase (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled recepto
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a col
CC   between the Swiss Institute of Bioinformatics and the EMBL out
CC   the European Bioinformatics Institute. There are no restrictio
CC   use by non-profit institutions as long as its content is li
CC   modified and this statement is not removed. Usage by and for
CC   entities requires a license agreement (See http://www.isb-sib.ch
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AJ245705; CAB53398.1; -
DR   EMBL; AY205086; AAP30960.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Phosphorylation; Lipoprotein; Palmitate.
FT   DOMAIN   1   37   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   38   63   1 (POTENTIAL).
FT   DOMAIN   64   72   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   73   93   2 (POTENTIAL).
FT   DOMAIN   94   118   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   119   140   3 (POTENTIAL).
FT   DOMAIN   141   163   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   164   183   4 (POTENTIAL).
FT   DOMAIN   184   191   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   192   211   5 (POTENTIAL).
FT   DOMAIN   212   240   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   241   266   6 (POTENTIAL).
FT   DOMAIN   267   279   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   280   300   7 (POTENTIAL).
FT   DOMAIN   301   317   CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD   29   29   N-LINKED (GLCNAC. . .) (POTENTIAL).

```
FT   METAL       333    333    IRON-SULFUR (4FE-4S) (BY SIMILARITY)
FT   METAL       336    336    IRON-SULFUR (4FE-4S) (BY SIMILARITY)
FT   METAL       340    340    IRON-SULFUR (4FE-4S) (BY SIMILARITY)
SQ   SEQUENCE    379 AA;  43014 MW;  9C257CCAD5547F5A CRC64;

Query Match
Best Local Similarity  100.0%;   Score 8;  DB 1;  Length 379;
Matches  8;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   217   LVDGVLAL  224
           ||||||||
Db    35   LVDGVLAL   42

RESULT 11
ZP3_MESAU
ID   ZP3_MESAU      STANDARD;      PRT;    422 AA.
AC   P23491;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Zona pellucida sperm-binding protein 3 precursor (Zona pellucida
DE   glycoprotein ZP3) (Sperm receptor) (Zona pellucida protein C).
GN   ZP3.
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=ovary;
RX   MEDLINE=91078540; PubMed=2257975;
RA   Kinloch R.A., Ruiz-Seller B., Wassarman P.M.;
RT   "Genomic organization and polypeptide primary structure of zona
RT   pellucida glycoprotein hZP3, the hamster sperm receptor.";
RL   Dev. Biol. 142:414-421(1990).
CC   -!- FUNCTION: Functions as a sperm-receptor. It is responsible fo
CC       sperm-adhesion to the zona pellucida, and may contribute to tl
CC       species-specificity of the insemination.
CC   -!- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN
CC       WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein. Extracellular
CC       matrix.
CC   -!- TISSUE SPECIFICITY: Oocytes.
CC   -!- DEVELOPMENTAL STAGE: GROWING OOCYTES.
CC   -!- PTM: Sulfated glycoprotein with O-linked oligosaccharides.
CC   -!- SIMILARITY: Contains 1 ZP domain.
CC   --------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a colla
CC   between the Swiss Institute of Bioinformatics and the EMBL outs
CC   the European Bioinformatics Institute. There are no restriction
CC   use by non-profit institutions as long as its content is in
CC   modified and this statement is not removed. Usage by and for co
CC   entities requires a license agreement (See http://www.isb-sib.ch/a
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------
DR   EMBL; M63629; AAA37079.1; -.
DR   InterPro; IPR001507; Endoglin/CD105.
DR   Pfam; PF00100; zona_pellucida; 1.
DR   PRINTS; PR00023; ZPELLUCIDA.
DR   SMART; SM00241; ZP; 1.
DR   PROSITE; PS00682; ZP_DOMAIN; 1.
KW   Glycoprotein; Signal; Sulfation; Sperm; Receptor; Transmembrane;
KW   Extracellular matrix.
FT   SIGNAL        1     22    POTENTIAL.
FT   CHAIN        23    422    ZONA PELLUCIDA SPERM-BINDING PROTEIN
FT   DOMAIN       23    386    EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    387    407    POTENTIAL.
FT   DOMAIN      408    422    CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       45    306    ZP.
FT   DOMAIN      119    158    PRO-RICH.
FT   DOMAIN      208    257    PRO-RICH.
```

Left column (partial, left margin cut off):

```
5    315    S-palmitoyl cysteine (Potential).
AA;  34699 MW;  6615D2146E1D247F CRC64;
rity  2.8%;  Score 8;  Pred. No.8;  DB 1;  Length 317;
nservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

IAA 144
IAA 167

STANDARD;      PRT;    379 AA.

el. 35, Created)
el. 35, Last sequence update)
el. 43, Last annotation update)
te dehydrogenase beta chain (EC 1.2.1.2).

jannaschii.
rchaeota; Methanococci; Methanococcales;
ccaceae; Methanocaldococcus.
0;

N.A.
DSM 2661 / ATCC 43067;
99; PubMed=868087;
te O., Olsen G.J., Zhou L., Fleischmann R.D.,
lake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
, Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
irkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
oghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
, Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
oberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
raser C.M., Smith H.O., Woese C.R., Venter J.C.;
me sequence of the methanogenic archaeon, Methanococcus

58-1073(1996).
ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
Binds 2 4Fe-4S clusters (Probable).
imer of alpha and beta chains (By similarity).
: The iron-sulfur centers are similar to those of
type 4Fe-4S ferredoxins.
: STRONG, TO THE BETA SUBUNIT OF M.THERMOAUTOTROPHICUM

T entry is copyright. It is produced through a collaboration
wiss Institute of Bioinformatics and the EMBL outstation -
ioinformatics Institute. There are no restrictions on its
rofit institutions as long as its content is in no way
his statement is not removed. Usage by and for commercial
res a license agreement (See http://www.isb-sib.ch/announce/
il to license@isb-sib.ch).

AAB97986.1; -.
64300.
-.
01450; 4Fe4S ferredoxin.
07525; FrhB_FdhB_C.
07516; FrhB_FdhB_N.
fer4; 1.
FrhB_FdhB_C; 1.
FrhB_FdhB_N; 1.
98; 4FE4S FERREDOXIN; 2.
protein; Oxidoreductase; NAD; Electron transport;
Fe-4S; Complete proteome.
0    280    IRON-SULFUR (4FE-4S) (BY SIMILARITY).
3    283    IRON-SULFUR (4FE-4S) (BY SIMILARITY).
6    286    IRON-SULFUR (4FE-4S) (BY SIMILARITY).
0    290    IRON-SULFUR (4FE-4S) (BY SIMILARITY).
0    330    IRON-SULFUR (4FE-4S) (BY SIMILARITY).
```

FT   146   146   N-LINKED (GLCNAC...) (POTENTIAL).
FT   271   271   N-LINKED (GLCNAC...) (POTENTIAL).
FT   302   302   N-LINKED (GLCNAC...) (POTENTIAL).
SQ   ...   22 AA; 45827 MW; D0F95BE7FF8E7E01 CRC64;

...larity   2.8%; Score 8; DB 1; Pred. No.10; Length 422;
...Conservative 0; Mismatches 0; Indels 0; Gaps 0;

...GLGLA 66
         |||||
...GLGLA 393

...STANDARD;   PRT;   576 AA.

(Rel. 32, Created)
(Rel. 32, Last sequence update)
(Rel. 41, Last annotation update)
...P-binding protein cydC.

...influenzae.
...oteobacteria; Gammaproteobacteria; Pasteurellales;
...ae; Haemophilus.
...27;

...N.A.
...KW20 / ATCC 51907;
...0630; PubMed=7542800;
R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
, Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
fitchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,

...a random sequencing and assembly of Haemophilus influenzae

...196-512(1995).
...SOMEHOW INVOLVED IN THE CYTOCHROME D BRANCH OF AEROBIC
...ON. SEEMS TO BE A COMPONENT OF A TRANSPORT SYSTEM
...ARITY).
...AR LOCATION: Integral membrane protein. Inner membrane
...a).
...TY: Belongs to the ABC transporter family. MsbA subfamily.

; AAC22811.1; -.
...E64186.

; AAC22811.1; -.
...R003593; AAA_ATPase.
...R001140; ABC_TM_transpt.
...R003439; ABC_transporter.
4; ABC_membrane; 1.
5; ABC_tran; 1.
...0006; ABC_transporter; 1.
82; AAA; 1.
...0929; ABC_TM1F; 1.
...0211; ABC_TRANSPORTER_1; 1.
...0893; ABC_TRANSPORTER_2; 1.
...Transport; Transmembrane; Inner membrane;
...eome.

FT   TRANSMEM   16   36   POTENTIAL.
FT   TRANSMEM   38   58   POTENTIAL.
FT   TRANSMEM   133   153   POTENTIAL.
FT   TRANSMEM   155   175   POTENTIAL.
FT   TRANSMEM   244   264   POTENTIAL.
FT   DOMAIN   281   301   POTENTIAL.
FT   DOMAIN   338   574   ABC TRANSPORTER.
FT   NP_BIND   372   379   ATP (POTENTIAL).
SQ   SEQUENCE   576 AA; 64831 MW; A9ACD8B9B294B1B3 CRC64;

Query Match   2.8%; Score 8; DB 1; Length 576;
Best_Local_Similarity 100.0%; Pred_No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0;

Qy   58 PLAIGLGL 65
         ||||||||
Db   159 PLAIGLGL 166

RESULT 13
GGT5_HUMAN
ID   GGT5_HUMAN   STANDARD;   PRT;   586 AA.
AC   P36269; Q96FC1; Q9UFM5;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Gamma-glutamyltransferase 5 precursor (EC 2.3.2.2) (Gamma-
DE   glutamyltransferase 5) (Gamma-glutamyltransferase-like activi
DE   (GGT-rel).
GN   GGTLA1 OR GGT5.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Placenta;
RX   MEDLINE=91296809; PubMed=1676842;
RA   Heisterkamp N., Rajpert-De Meyts E., Uribe L., Forman H.J.,
RA   Groffen J.;
RT   "Identification of a human gamma-glutamyl cleaving enzyme relate
RT   but distinct from, gamma-glutamyl transpeptidase.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:6303-6307(1991).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   TISSUE=Fetal kidney;
RA   Blum H., Bauersachs S., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20057165; PubMed=10591208;
RA   Dunham I., Hunt A.R., Collins J.E., Bruskiewich R., Beare D.M.,
RA   Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA   Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA   Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA   Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark (
RA   Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
RA   Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson]
RA   Dhami P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A
RA   Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
RA   Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hal
RA   Hall R.E., Hall-Tamlyn G., Heathcott R.W., Ho S., Holmes S.,
RA   Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
RA   Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,
RA   Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
RA   Mcclay J., Mclaren S., McMurray A.A., Milne S.A., Mortimore B.J.
RA   Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C
RA   Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross
RA   Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
RA   Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.
RA   Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
RA   Williams S.A., Williamson H., Wilmer T.E., Wilming]
RA   Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimi

RA   Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
RA   Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
RA   ... F., Chu L., Crabtree J., Deschamps J., Do T.,
RA   ng F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
RA   is S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
RA   ., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
RA   Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
RA   g G., Chissoe S., Murray J., Miller N., Minx P.,
RA   nnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
RA   Z., Fulton L., Goela D., Graves T., Hawkins J.,
RA   p K., Latreille P., Layman D., Ozersky P., Rohling T.,
RA   lker C., Wamsley A., Wohldmann P., Pepin K., Nelson J.,
RA   ll J.A., Hillier L.W., Mardis E., Waterston R.,
RA   anuel B.S., Shaikh T., Kurahashi H., Saitta S.,
RA   McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
RA   Kim U.J., Shizuya H., Simon M.I., Dumanski J.P.,
RA   edra D., Seroussi E., Fransson I., Tapia I., Bruder C.E.,
RA   Wilkinson P., Bodenteich A., Hartman K., Hu X.,
RA   ne L., Tilahun Y., Wright H.;
RL   ence of human chromosome 22.";
RL   9-495(1999).

RN   N.A. (ISOFORM 1).
RX   257; PubMed=12477932;
RA   J., Feingold E.A., Grouse L.H., Derge J.G.,
RA   ., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   quellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   cEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   ton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   adan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   ., Touchman J.W., Green E.D., Dickson M.C.,
RA   ., Grimwood J., Schmutz J., Myers R.M.,
RA   .S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   chein J.E., Jones S.J.M., Marra M.A.;
RL   nd initial analysis of more than 15,000 full-length
RL   cad. Sci. U.S.A. 99:16899-16903(2002).

CC   se cDNA sequences.";
CC   Cleaves the gamma-glutamyl peptide bond of glutathione
CC   s, but maybe not glutathione itself. Converts
CC   ne C4 (LTC4) to leukotriene D4 (LTD4).
CC   ACTIVITY: (5-L-glutamyl)-peptide + an amino acid =
CC   5-L-glutamyl-amino acid.
CC   leukotriene metabolism; second step.
CC   Heterodimer composed of the light and heavy chains.
CC   e site is located in the light chain (By similarity).
CC   Y: Belongs to the gamma-glutamyltransferase family.
CC   AR LOCATION: Type II membrane protein (By similarity).
CC   VE PRODUCTS:
CC   ernative splicing; Named isoforms=2;

CC   36269-1; Sequence=Displayed;
CC   36269-2; Sequence=VSP_008146;
CC   experimental confirmation available;
CC   Ref.2 sequence differs from that shown due to a
CC   t in position 446.

CC   DT entry is copyright. It is produced through a collaboration
CC   Swiss Institute of Bioinformatics and the EMBL outstation -
CC   Bioinformatics Institute. There are no restrictions on its
CC   profit institutions as long as its content is in no way
CC   ires a license agreement (See http://www.isb-sib.ch/announce/
CC   this statement is not removed. Usage by and for commercial
CC   ail to license@isb-sib.ch).

CC   --
DR   EMBL; M64099; AAA58503.1; -.
DR   EMBL; AL117414; CAB55910.1; ALT_FRAME.
DR   EMBL; AP000354; -; NOT_ANNOTATED_CDS.
DR   EMBL; BC011362; AAH11362.1; -.
DR   PIR; A41125; A41125.
DR   PIR; T17220; T17220.
DR   MEROPS; T03.002; -.
DR   Genew; HGNC:4260; GGTLA1.
DR   MIM; 137168; -.
DR   GO; GO:0016021; C:integral to membrane; TAS.
DR   GO; GO:0003840; F:gamma-glutamyltransferase activity; TAS.
DR   GO; GO:0003840; P:amino acid metabolism; TAS.
DR   GO; GO:0006749; P:glutathione metabolism; TAS.
DR   InterPro; IPR000101; Peptidase T3.
DR   Pfam; PF01019; G_glu_transpept; 1.
DR   PRINTS; PR01210; GGTRANSPASE.
DR   PROSITE; PS00462; G_GLU_TRANSPEPTIDASE; 1.
KW   Glutathione biosynthesis; Leukotriene biosynthesis; Transferase;
KW   Acyltransferase; Signal-anchor; Transmembrane; Zymogen; Glycoprot
KW   Alternative splicing;
FT   CHAIN   1   387   GAMMA-GLUTAMYLTRANSFERASE 5 HEAVY
FT   CHAIN (BY SIMILARITY).
FT   CHAIN   388   586   GAMMA-GLUTAMYLTRANSFERASE 5 LIGHT
FT   CHAIN (BY SIMILARITY).
FT   DOMAIN   1   8   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   9   29   SIGNAL-ANCHOR (TYPE-II MEMBRANE PROT
FT   (POTENTIAL).
FT   DOMAIN   30   586   EXTRACELLULAR (POTENTIAL).
FT   CARBOHYD   98   98   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   204   204   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   303   303   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   347   347   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   535   535   N-LINKED (GLCNAC...) (POTENTIAL).
FT   VARSPLIC   101   132   Missing (in isoform 2).
FT   /FTId=VSP_008146.
FT   CONFLICT   330   330   R -> K (IN REF. 2 AND 4).
FT   CONFLICT   408   408   N -> Y (IN REF. 2).
FT   CONFLICT   437   437   W -> R (IN REF. 2 AND 4).
SQ   SEQUENCE   586 AA; 62319 MW; 1BE543CB0934B16B CRC64;

Query Match   2.8%; Score 8; DB 1; Length 586;
Best Local Similarity 100.0%; Pred. No. 14;
Matches   8; Conservative   0; Mismatches   0; Indels   0; G

Qy   61 LGLGLALA 68
              ||||||||
Db   14 LGLGLALA 21

RESULT 14
ID   HSP3_OCTVU   STANDARD;   PRT;   24 AA.
AC   P83215;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Sperm protamine P3 (Po3) (Fragment).
OS   Octopus vulgaris (Octopus).
OC   Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoide
OC   Octopodiformes; Octopoda; Incirrata; Octopodidae; Octopus.
OX   NCBI_TaxID=6645;
RN   [1]
RP   SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC   TISSUE=Sperm;
RA   Gimenez-Bonafe P., Ribes E., Buesa C., Sautiere P., Kouach M.,
RA   Ausio J., Kasinsky H.E., Chiva M.;
RT   "Chromatin remodelling and protamines during spermiogenesis of Oc
RT   vulgaris (Cephalopoda).";
RL   J. Exp. Zool. 0:0-0(2001).
CC   -!- FUNCTION: Protamines substitute for histones in the chromatin
CC   sperm during the haploid phase of spermatogenesis. They compa

Db   5 RRRGRRG 11

RESULT 16
HSP1_SAGIM
ID   HSP1_SAGIM   STANDARD;   PRT;   49 AA.
AC   P24714;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Sperm protamine P1 (Cysteine-rich protamine).
GN   PRM1.
OS   Saguinus imperator (Emperor tamarin).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
OC   Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Sagui
OX   NCBI_TaxID=9491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=92051332; PubMed=1840669;
RA   Queralt R., Oliva R.;
RT   "Protamine 1 gene sequence from the primate Saguinus imperator
RT   isolated with PCR using consensus oligonucleotides.";
RL   Nucleic Acids Res. 19:5786-5786(1991).
CC   -!- FUNCTION: Protamines substitute for histones in the chromati
CC   sperm during the haploid phase of spermatogenesis. They comp
CC   sperm DNA into a highly condensed, stable and inactive compl
CC   -!- SUBUNIT: Cross-linked by interchain disulfide bonds around t
CC   DNA-helix (By similarity).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- TISSUE SPECIFICITY: Testis.
CC   -!- SIMILARITY: Belongs to the protamine P1 family.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a col
CC   between the Swiss Institute of Bioinformatics and the EMBL ou
CC   the European Bioinformatics Institute. There are no restrictio
CC   use by non-profit institutions as long as its content is i
CC   modified and this statement is not removed. Usage by and for
CC   entities requires a license agreement (See http://www.isb-sib.ch
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; X61678; CAA43853.1; -.
DR   PIR; S22582; S22582.
DR   InterPro; IPR000221; Protamine_P1.
DR   Pfam; PF00260; protamine P1; 1.
DR   PROSITE; PS00048; PROTAMINE_P1; 1.
KW   Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-bindi
KW   Testis; DNA condensation; Nuclear protein.
FT   INIT_MET   0   0
SQ   SEQUENCE   49 AA;   6545 MW;   8389C403F5B207F6 CRC64;

Query Match   2.5%;
Best Local Similarity   100.0%;   Score 7;   DB 1;   Length 49;
Matches   7;   Conservative   0;   Mismatches   0;   Indels   0;

Qy   42 QRRRGRR 48
         |||||||
Db   17 QRRRGRR 23

RESULT 17
HSP1_DIDMA
ID   HSP1_DIDMA   STANDARD;   PRT;   57 AA.
AC   P35305;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Sperm protamine P1.
GN   PRM1.
OS   Didelphis marsupialis virginiana (North American opossum), and
OS   Monodelphis domestica (Short-tailed grey opossum).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
OC   Mammalia; Metatheria; Didelphimorphia; Didelphidae; Didelphis.

A into a highly condensed, stable and inactive

AR LOCATION: Nuclear.
PECIFICITY: Testis.
CTROMETRY: MW=4389; METHOD=Electrospray.
86; C:nucleosome; NAS.
34; C:nucleus; NAS.
77; F:DNA binding; NAS.
01; P:chromosome organization and biogenesis (sen. ..; NAS.
76; P:mitotic chromosome condensation; NAS.
34; P:nucleosome assembly; NAS.
83; P:spermatogenesis; NAS.
protein; Nucleosome core; Spermatogenesis;
DNA condensation; Nuclear protein.
1   16   POLY-ARG.
24   24
4 AA;   3381 MW;   308E90ED9D2C9C9C CRC64;

larity 100.0%; Pred. No.7.7;   Length 24;
Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

GRRG 49
||||
GRRG 17

STANDARD;   PRT;   30 AA.

(Rel. 41, Created)
(Rel. 41, Last sequence update)
(Rel. 41, Last annotation update)
ine P5 (Po5).
aris (Octopus).
etazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
es; Octopoda; Incirrata; Octopodidae; Octopus.
645;

NCTION, AND MASS SPECTROMETRY.

;
fe P., Ribes E., Buesa C., Sautiere P., Kouach M.,
sinsky H.E., Chiva M.;
emodelling and protamines during spermiogenesis of Octopus
phalopoda).";
: 0:0-0(2001).
: Protamines substitute for histones in the chromatin of
ring the haploid phase of spermatogenesis. They compact
A into a highly condensed, stable and inactive

AR LOCATION: Nuclear.
PECIFICITY: Testis.
CTROMETRY: MW=3941; METHOD=Electrospray.
86; C:nucleosome; NAS.
34; C:nucleus; NAS.
77; F:DNA binding; NAS.
01; P:chromosome organization and biogenesis (sen. ..; NAS.
76; P:mitotic chromosome condensation; NAS.
34; P:nucleosome assembly; NAS.
83; P:spermatogenesis; NAS.
protein; Nucleosome core; Spermatogenesis;
DNA condensation; Nuclear protein.
2   15   POLY-ARG.
17   26   POLY-ARG.
0 AA;   3943 MW;   14F1BC7E4D277049 CRC64;

larity 2.5%; Score 7; DB 1; Length 30;
Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

GRRG 49
||||

57, 13616;
N.A.
...supialis;
500; PubMed=8344286;
Nishikawa S., Connor W., Dixon G.H.;
...ion of a marsupial sperm protamine gene and its
...rom the North American opossum (Didelphis
...;
...em. 215:63-72(1993).
N.A.
...astica;
...951; PubMed=7700877;
Grajewski C., Westerman M., Winkfein R.J., Dixon G.H.;
...ylogeny and evolution of marsupial protamine P1 genes.";
...Lond., B, Biol. Sci. 259:7-14(1995).
Protamines substitute for histones in the chromatin of
...ng the haploid phase of spermatogenesis. They compact
...into a highly condensed, stable and inactive complex.
...AR LOCATION: Nuclear.
...CIFICITY: Testis.
...: Belongs to the protamine P1 family.

AAA02812.1; -.
CAA52193.1; -.
AAA74612.1; -.
...34045.
000221; Protamine_P1.
...protamine_P1; 1.
...048; PROTAMINE P1; 1.
...otein; Nucleosome core; Spermatogenesis; DNA-binding;
...ondensation; Nuclear protein.
0   0   BY SIMILARITY.
...AA; 7810 MW; 28371SB280214E52 CRC64;

...rity   2.5%;   Score 7;   DB 1;   Length 57;
100.0%;   Pred.No.16;
...nservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

RG 49
||
RG 40

STANDARD;   PRT;   115 AA.
...Rel. 40, (Created)
...Rel. 40, Last sequence update)
...Rel. 42, Last annotation update)
...nd protein Acp62F precursor.
...262.
...lanogaster (Fruit fly).
...azoa; Arthropoda; Hexapoda; Insecta; Pterygota;
...opterygota; Diptera; Brachycera; Muscomorpha;
...Drosophilidae; Drosophila.
...37;
N.A., FUNCTION, AND TISSUE SPECIFICITY.
...S; TISSUE=Male accessory gland;
...20; PubMed=9474779;
Harada H.A., Bertram M.J., Stelick T.J., Kraus K.W.,
...ng Y.O., Neubaum D.M., Park M., Tram U.K.,

RT   "New genes for male accessory gland proteins in Drosophila
RT   melanogaster";
RL   Insect Biochem. Mol. Biol. 27:825-834(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B., Dun
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischman
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [3]
RP   SEQUENCE OF 7-111 FROM N.A.
RC   STRAIN=ZIM62H-12C, ZIM62H-16C, ZIM62H-28C, ZIM62H-30C, ZIM62H-34C
RC   ZIM62I-5C, ZIM62H-10C, ZIM62I-17C, ZIM62I-18C, and ZIM62I-53C;
RX   MEDLINE=20556153; PubMed=11102381;
RA   Begun D.J., Whitley P., Todd B.L., Waldrip-Dail H.M., Clark A.G.;
RT   "Molecular population genetics of male accessory gland proteins i
RT   Drosophila.";
RL   Genetics 156:1879-1888(2000).
CC   -!- FUNCTION: RESPONSIBLE FOR PHYSIOLOGICAL AND BEHAVIORAL CHANGE
CC   MATED FEMALE FLIES. MAY CONTRIBUTE TO THE TOXICITY OF SEMINAL
CC   FLUID AND THE DECREASED LIFE-SPAN OF MATED FEMALES. MAY ALSO
CC   AFFECT NEUROMUSCULAR EVENTS AFTER MATING CONCERNING SPERM STO
CC   AND EGG RELEASE.
CC   -!- SUBCELLULAR LOCATION: Secreted (Probable).
CC   -!- TISSUE SPECIFICITY: Seminal fluid.
CC   -!- SIMILARITY: SOME, TO P.NIGRIVENTER TX2-6.
DR   EMBL; U85763; AAB96387.1; -.
DR   EMBL; AB003475; AAF47683.1; -.

RX   MEDLINE=97251357; PubMed=9097039;
RA   Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA   Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
RA   Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,
RA   Motomura K., Nakade S., Nakamura Y., Nashimoto H., Nishio Y.,
RA   Oshima T., Saito N., Sampei G., Seki Y., Sivasundaram S.,
RA   Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
RA   Yamamoto Y., Horiuchi T.;
RT   "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 28.0-40.1 min region on the linkage map.";
RL   DNA Res. 3:363-377(1996).
CC   This SWISS-PROT entry is copyright. It is produced through a col
CC   between the Swiss Institute of Bioinformatics and the EMBL out
CC   the European Bioinformatics Institute. There are no restriction
CC   use by non-profit institutions as long as its content is i
CC   modified and this statement is not removed. Usage by and for (
CC   entities requires a license agreement (See http://www.isb-sib.ch.
CC   or send an email to license@isb-sib.ch).
DR   EMBL; J01734; -; NOT ANNOTATED_CDS.
DR   EMBL; U70214; AAB08680.1; -.
DR   EMBL; D83536; -; NOT ANNOTATED_CDS.
DR   EMBL; D90771; BAA14925.1; -.
DR   EMBL; D90772; BAA14935.1; -.
DR   EMBL; D90831; BAA15715.1; -.
DR   EMBL; D90841; BAA15872.1; -.
DR   EMBL; D90847; BAA15958.1; -.
DR   EMBL; D90848; BAA15963.1; -.
DR   PIR; B91483; IEECSB.
KW   Hypothetical protein; Transposable element.
SQ   SEQUENCE 118 AA; 12270 MW; 348014FAC765058E CRC64;

Query Match   2.5%; Score 7; DB 1; Length 118;
  Best Local Similarity   100.0%; Pred. No. 31;
  Matches   7; Conservative   0; Mismatches   0; Indels   0; (

Qy   23 DGGAVRQ 29
         |||||||
Db   107 DGGAVRQ 113

RESULT 20
IL13_MOUSE
ID   IL13_MOUSE   STANDARD;   PRT;   131 AA.
AC   P20109;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Interleukin-13 precursor (IL-13) (T-cell activation protein P600
GN   IL13 OR IL-13.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mu
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89093958; PubMed=2521353;
RA   Brown K.D., Zurawski S.M., Mosmann T.R., Zurawski G.;
RT   "A family of small inducible proteins secreted by leukocytes are
RT   members of a new superfamily that includes leukocyte and
RT   fibroblast-derived inflammatory agents, growth factors, and
RT   indicators of various activation processes.";
RL   J. Immunol. 142:679-687(1989).
CC   -!- FUNCTION: CYTOKINE. INHIBITS INFLAMMATORY CYTOKINE PRODUCTION
CC   SYNERGIZES WITH IL2 IN REGULATING INTERFERON-GAMMA SYNTHESIS
CC   MAY BE CRITICAL IN REGULATING INFLAMMATORY AND IMMUNE RESPONS
CC   (BY SIMILARITY)
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the IL-4 / IL-13 family.
CC   This SWISS-PROT entry is copyright. It is produced through a col
CC   between the Swiss Institute of Bioinformatics and the EMBL out

---

DR   8;   AAG35367.1; -
DR   9;   AAG35368.1; -
DR   0;   AAG35369.1; -
DR   1;   AAG35370.1; -
DR   2;   AAG35371.1; -
DR   3;   AAG35372.1; -
DR   4;   AAG35373.1; -
DR   5;   AAG35374.1; -
DR   6;   AAG35375.1; -
DR   7;   AAG35376.1; -.
DR   0020509; Acp62F.
FT   7; F:serine protease inhibitor activity; IDA.
FT   0; P:determination of adult life span; NAS.
FT   2002919; TIL_Cysrich.
FT   ; TIL; 1.
FT   nal.
PT   1   24   POTENTIAL.
            ACCESSORY GLAND PROTEIN ACP62F.
PT   25   115   TIL.
PT   34   88
SQ   5 AA; 12570 MW; 4326A6F6C32291D CRC64;

arity   2.5%;   Score 7; DB 1; Length 115;
             100.0%; Pred. No. 30;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   LLL 74
         |||
Dt   LLL 17

ID   STANDARD;   PRT;   118 AA.
DT   (Rel. 01, Created)
DT   (Rel. 01, Last sequence update)
DT   (Rel. 41, Last annotation update)
DE   ment IS5 very hypothetical 12 kDa protein.
OS   oli.
OC   oteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   aceae; Escherichia.
RN   [1] N.A.
     653; PubMed=6269959;
     ahn M.;
RT   de sequence of IS5 from Escherichia coli.";
RL   74(1981).
RN   [1] N.A.
     652; PubMed=6269958;
     van Bree M.P.;
RT   de sequence and protein-coding capability of the
RT   element IS5.";
RL   63(1981).
RN   [1] N.A.
RA   Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,
RA   erspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
RA   Lew H., Lin D., Namath A., Oefner P., Roberts D.;
RL   EP-1996) to the EMBL/GenBank/DDBJ databases.
RN   [1] N.A.
RA   Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
RA   Inokuchi H., Miki T., Hatada E., Fukuda R.,
RA   Mizuno T., Makino K., Nakata A., Yura T., Sampei G.,
RL   ; EB-1996) to the EMBL/GenBank/DDBJ databases.
RN   [1] N.A.

```
DR   HSSP; P35225; 3ITR.
DR   InterPro; IPR003634; Interleukin_13.
DR   InterPro; IPR001325; Interleukin_4_13.
DR   Pfam; PF03487; Interleukin_13; 1.
DR   ProDom; PD015987; Interleukin_13; 1.
DR   SMART; SM00190; IL4_13; 1.
DR   PROSITE; PS00838; INTERLEUKIN_4_13; 1.
KW   Cytokine; Glycoprotein; Signal.
FT   SIGNAL      1     21       BY SIMILARITY.
FT   CHAIN      22    131       INTERLEUKIN-13.
FT   DISULFID   52     80       BY SIMILARITY.
FT   DISULFID   68     94       BY SIMILARITY.
FT   CARBOHYD   42     42       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   53     53       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   76     76       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD  121    121       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE 131 AA;  14093 MW;  E5008CAB8DE8C201 CRC64;

Query Match             2.5%;  Score 7;  DB 1;  Length 131;
Best Local Similarity 100.0%;  Pred. No. 34;
Matches   7;  Conservative   0;  Mismatches   0;  Indels   0;  Ga...

Qy   65 LALACLG 71
        |||||||
Db    9 LALACLG 15


RESULT 22
YK01_PYRHO
ID   YK01_PYRHO   STANDARD;   PRT;   147 AA.
AC   O57781;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein PH2001.
GN   PH2001.
OS   Pyrococcus horikoshii.
OC   Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcac...
OC   Pyrococcus.
OX   NCBI_TaxID=53953;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OT3;
RX   MEDLINE=98344137; PubMed=9679194;
RA   Kawarabayasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA   Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai
RA   Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y
RA   Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguch
RA   Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA   Masuchi Y., Shizuya H., Kikuchi H.;
RT   "Complete sequence and gene organization of the genome of a hyper
RT   thermophilic archaebacterium, Pyrococcus horikoshii OT3.";
RL   DNA Res. 5:55-76(1998).
```

```
DR   EMBL; AP000001; BAA31940.1; -.
DR   EMBL; AP000007; BAA31940.1; JOINED.
DR   EMBL; AP000007; BAA31943.1; -.
DR   EMBL; AP000001; BAA31943.1; JOINED.
DR   PIR; A71217; A71217.
KW   Hypothetical protein; Transmembrane; Complete proteome.
FT   TRANSMEM   41     61       POTENTIAL.
FT   TRANSMEM   67     87       POTENTIAL.
SQ   SEQUENCE 147 AA;  15324 MW;  247ED12FCEF265B9 CRC64;

Query Match             2.5%;  Score 7;  DB 1;  Length 147;
```

```
AAA40149.1; -.
30552.
3ITR.
; I113.
03634; Interleukin_13.
01325; Interleukin_4_13.
     Interleukin_13; 1.
87; Interleukin_13; 1.
0; IL4_13; 1.
38; INTERLEUKIN_4_13; 1.
oprotein; Signal.
1     21       BY SIMILARITY.
2    131       INTERLEUKIN-13.
1     79       BY SIMILARITY.
1     93       BY SIMILARITY.
2     42       N-LINKED (GLCNAC...) (POTENTIAL).
2     52       N-LINKED (GLCNAC...) (POTENTIAL).
5     75       N-LINKED (GLCNAC...) (POTENTIAL).
AA;  14107 MW;  954F93F105713FED CRC64;

rity       2.5%;  Score 7;  DB 1;  Length 131;
onservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

LG 71
||
LG 15

STANDARD;   PRT;   131 AA.

el. 32, Created)
el. 32, Last sequence update)
el. 36, Last annotation update)
 precursor (IL-13) (T-cell activation protein P600).

cus (Rat).
azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
neria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
N.A.
-Dawley; TISSUE=Kidney cortex;
38; PubMed=7916615;
ruet E.N.;
t interleukin-13 (IL-13) cDNA and analysis of IL-13
n in experimental glomerulonephritis.";
ys. Res. Commun. 197:612-618(1993).
CYTOKINE. INHIBITS INFLAMMATORY CYTOKINE PRODUCTION.
 WITH IL2 IN REGULATING INTERFERON-GAMMA SYNTHESIS.
TICAL IN REGULATING INFLAMMATORY AND IMMUNE RESPONSES
ARITY].
AR LOCATION: Secreted.
: Belongs to the IL-4 / IL-13 family.
```

```
AAA16478.1; -.
52290.
```

...arity 100.0%; Pred. No. 38;
...onservative   0; Mismatches   0; Indels   0; Gaps   0;

...GLL 73
    |||
...GLL 51

        STANDARD;   PRT;   150 AA.

Rel. 41, Created)
Rel. 41, Last sequence update)
Rel. 41, Last annotation update)
...eta (LT-beta) (Tumor necrosis factor C) (TNF-C) (Tumor
...or ligand superfamily member 3) (Fragment).
OR TNFC.
...g).
...tazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
...heria; Cetartiodactyla; Suina; Suidae; Sus.
...23;
...1 N.A.
...white; TISSUE=Fibroblast;
...615; PubMed=1169259;
...ogel-Gaillard C., Cattolico L., Duprat S., Vaiman M.,
...the swine major histocompatibility complex region
...l non-classical class I genes.";
...ns 57:55-65(2001).
...   Cytokine that binds to LTBR/TNFRSF3. May play a specific
...mmune response regulation. Provides the membrane anchor
...ttachment of the heterotrimeric complex to the cell
...Heterotrimer of either two LTB and one LTA subunits or
...valent) two LTA and one LTB subunits.
...AR LOCATION: Type II membrane protein (By similarity).
...Y: Belongs to the tumor necrosis factor family.
...OT entry is copyright. It is produced through a collaboration
Swiss Institute of Bioinformatics and the EMBL outstation -
Bioinformatics Institute. There are no restrictions on its
profit institutions as long as its content is in no way
this statement is not removed. Usage by and for commercial
ires a license agreement (See http://www.isb-sib.ch/announce/
ail to license@isb-sib.ch).
4; CAB63851.1; -.
006053; TNF_abc.
006052; TNF_family.
008983; TNF_like.
003636; TNF_subf.
; TNF; 1.
34; TNECROSISFCT.
012; TNF_subf; 1.
7; TNF; 1.
251; TNF_1; 1.
049; TNF_2; 1.
...coprotein.
1   1
28   128   N-LINKED (GLCNAC...) (POTENTIAL).
0 AA; 16423 MW; FE5C4CC657658B48 CRC64;

...arity   2.5%; Score 7; DB 1; Length 150;
...100.0%; Pred. No. 38;
...onservative   0; Mismatches   0; Indels   0; Gaps   0;

LYC 200
    |||
LYC 43

RESULT 24
RA05_ORYSA
ID   RA05_ORYSA   STANDARD;   PRT;   157 AA.
AC   Q01881;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Seed allergenic protein RA5 precursor.
GN   RA5.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Seed;
RX   MEDLINE=93144699; PubMed=767876S;
RA   Adachi T., Izumi H., Yamada H., Tanaka K., Takeuchi S.,
RA   Nakamura R., Matsuda T.;
RT   "Gene structure and expression of rice seed allergenic proteins
RT   belonging to the alpha-amylase/trypsin inhibitor family.";
RL   Plant Mol. Biol. 21:239-248(1993).
CC   -!- PTM: Five disulfide bonds are present (By similarity).
CC   -!- ALLERGEN: Causes an allergic reaction in human.
CC   -!- SIMILARITY: Belongs to the cereal trypsin/alpha-amylase inhibitor
CC   family.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restriction
CC   use by non-profit institutions as long as its content is in
CC   modified and this statement is not removed. Usage by and for c
CC   entities requires a license agreement (See http://www.isb-sib.ch/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; D11430; BAA01996.1; -.
DR   PIR; S31078; S31078.
DR   HSSP; P01085; 1HSS.
DR   Gramene; Q01881; -.
DR   InterPro; IPR003612; AAI.
DR   InterPro; IPR006106; Amylase_inhib.
DR   InterPro; IPR006105; Try/amyl_inhib.
DR   Pfam; PF00234; tryp_alpha_amyl; 1.
DR   PRINTS; PR00808; AMLASEINHBTR.
DR   SMART; SM00499; AAI; 1.
DR   PROSITE; PS00426; CEREAL_TRYP_AMYL_INH; 1.
KW   Allergen; Multigene family; Signal.
FT   SIGNAL   1   26   POTENTIAL.
FT   CHAIN   27   157   SEED ALLERGENIC PROTEIN RA5.
SQ   SEQUENCE   157 AA; 17118 MW; C8A5495FBFB399E6 CRC64;

Query Match   2.5%; Score 7; DB 1; Length 157;
Best Local Similarity 100.0%; Pred. No. 40;
Matches   7; Conservative   0; Mismatches   0; Indels   0; G

Qy   72 LLLAVVS 78
              |||||||
Db   11 LLLAVVS 17

RESULT 25
HLPA_ECOLI
ID   HLPA_ECOLI   STANDARD;   PRT;   161 AA.
AC   P11457;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Histone-like protein HLP-1 precursor (DNA-binding 17 kDa protein)
GN   HLPA OR SKP OR OMPH OR B0178 OR C0215 OR Z0190 OR ECS0180 OR SF01
GN   OR S0171.
OS   Escherichia coli.
OS   Escherichia coli O6.

li O157:H7, and
eri.

eobacteria; Gammaproteobacteria; Enterobacteriales;
ceae; Escherichia.
, 217992, 83334, 623;

N.A., AND SEQUENCE OF 21-40.

35; PubMed=2843433;
pe K.;
sequencing of the gene for the DNA-binding 17K protein
coli.";
24 (1988).

N.A.
; STRAIN=K12 / MG1655;
17; PubMed=9278503;
Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
ado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
ris N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
.;
genome sequence of Escherichia coli K-12.";
53-1474 (1997).

N.A.
; STRAIN=K12 / W3110;
ori H., Murayama N., Kataoka K., Yano M., Itoh T.,
nokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,
ino K., Nakata A., Yura T., Sampei G., Mizobuchi K.;
quencing of the Escherichia coli genome: analysis of the
(189,987 - 281,416bp) region.";
-1996) to the EMBL/GenBank/DDBJ databases.

N.A.
ncan M., Allen E., Araujo R., Aparicio A., Chung E.,
rland V., Plunkett G. III, Redford P., Roesch P.,
les E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
ose D.J., Zhou S., Schwartz D.C., Perna N.T.,
Donnenberg M.S., Blattner F.R.;
aic structure revealed by the complete genome sequence
ic Escherichia coli.";
ad. Sci. U.S.A. 99:17020-17024(2002).

N.A.
; STRAIN=O6:H1 / CFT073 / ATCC 700928;

N.A.
; STRAIN=O157:H7 / EDL933 / ATCC 700927;
35; PubMed=11206551;
unkett G. III, Burland V., Mau B., Glasner J.D.,
hew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
kett J., Klink S., Boutin A., Shao Y., Miller L.,
Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
antharaman T.S., Lin J., Yen G., Schwartz D.C.,
attner F.R.;
ce of enterohaemorrhagic Escherichia coli O157:H7.";
-533 (2001).

N.A.
; STRAIN=O157:H7 / RIMD 0509952;
31; PubMed=11258796;
kino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
subo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
ni H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
ba T., Hattori M., Shinagawa H.;
ome sequence of enterohemorrhagic Escherichia coli
nomic comparison with a laboratory strain K-12.";

RL   DNA Res. 8:11-22(2001).
RN   [8]
RP   SEQUENCE OF 72-161 FROM N.A.
RC   SPECIES=E.coli; STRAIN=K12 / MG1655;
RX   MEDLINE=91100302; PubMed=1987124;
RA   Dicker I.B., Seetharam S.R.;
RT   "Cloning and nucleotide sequence of the firA gene and the firA200
RT   allele from Escherichia coli.";
RL   J. Bacteriol. 173:334-344(1991).
RN   [9]
RP   SEQUENCE OF 21-32.
RC   SPECIES=E.coli; STRAIN=K12 / EMG2;
RX   MEDLINE=97443975; PubMed=9298646;
RA   Link A.J., Robison K., Church G.M.;
RT   "Comparing the predicted and observed properties of proteins enco
RT   in the genome of Escherichia coli K-12.";
RL   Electrophoresis 18:1259-1313(1997).
RN   [10]
RP   SEQUENCE FROM N.A.
RC   SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
RX   MEDLINE=22272406; PubMed=12384590;
RA   Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA   Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong
RA   Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA   Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA   Yu J.;
RT   "Genome sequence of Shigella flexneri 2a: insights into pathogeni
RT   through comparison with genomes of Escherichia coli K12 and O157.
RL   Nucleic Acids Res. 30:4432-4441(2002).
RN   [11]
RP   SEQUENCE FROM N.A.
RC   SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX   MEDLINE=22590274; PubMed=12704152;
RA   Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA   Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA   Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA   Schwartz D.C., Blattner F.R.;
RT   "Complete genome sequence and comparative genomics of Shigella
RT   flexneri serotype 2a strain 2457T.";
RL   Infect. Immun. 71:2775-2786(2003).
RN   [12]
RP   SIMILARITY TO S.TYPHIMURIUM OMPH.
RC   SPECIES=E.coli;
RX   MEDLINE=90201355; PubMed=2318304;
RA   Hirvas L., Coleman J., Koski P., Vaara M.;
RT   "Bacterial 'histone-like protein I' (HLP-I) is an outer membrane
RT   constituent.";
RL   FEBS Lett. 262:123-126(1990).
CC   -!- SUBUNIT: Homotetramer.
CC   -!- SUBCELLULAR LOCATION: EITHER IN THE NUCLEOID (CHROMATIN) OR I
CC   THE OUTER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE OMPH/HLPA FAMILY.
DR   EMBL; M21118; AAA24630.1; -.
DR   EMBL; AE000127; AAC73289.1; -.
DR   EMBL; D83536; BAA77853.1; -.
DR   EMBL; U70214; AAB08607.1; -.
DR   EMBL; AE016755; AAN78707.1; -.
DR   EMBL; AE005194; AAG54480.1; -.
DR   EMBL; AP002550; BAB33603.1; -.
DR   EMBL; X54797; CAA38567.1; -.
DR   EMBL; X75465; CAA53207.1; -.
DR   EMBL; AE015054; AAN41830.1; -.
DR   EMBL; AE016978; AAP15711.1; -.
DR   PIR; D85502; D85502.

D90651.
DNEC17.
455; hlpA.
005632; OmpH.
; OmpH; 1.
Outer membrane; Signal; Complete proteome.
1   20
21   161   HISTONE-LIKE PROTEIN HLP-1.
1 AA; 17688 MW; 2A966BBD83F3E675 CRC64;

arity 100.0%; Score 7; DB 1; Length 161;
onservative 0; Mismatches 0; Indels 0; Gaps 0;

ALA 68
|||
ALA 15

STANDARD;   PRT;   170 AA.

(Rel. 28, Created)
(Rel. 28, Last sequence update)
(Rel. 41, Last annotation update)
protein D2007.4 in chromosome III.

s elegans.
etazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
Peloderinae; Caenorhabditis.
39;

N.A.
1 N2;
718; PubMed=7906398;
nscough R., Anderson K., Baynes C., Berks M.,
Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
ear S., Du Z., Durbin R., Favello A., Fraser A.,
rdner A., Green P., Hawkins T., Hillier L., Jier M.,
Jones M., Kershaw J., Kirsten J., Laister N.,
Lightning J., Lloyd C., Mortimore B. O'Callaghan M.,
Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
don N., Smith A., Smith M., Sonnhammer E., Staden R.,
hierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
Watson A., Weinstock L., Wilkinson-Sproat J.,

ontiguous nucleotide sequence from chromosome III of C.

-38(1994).

OT entry is copyright. It is produced through a collaboration
Swiss Institute of Bioinformatics and the EMBL outstation -
Bioinformatics Institute. There are no restrictions on its
profit institutions as long as its content is in no way
this statement is not removed. Usage by and for commercial
ires a license agreement (See http://www.isb-sib.ch/announce/
ail to license@isb-sib.ch).

AAA27999.1; -.
S44789.
07.4; CE00129.
protein.
70 AA; 19396 MW; 22301D7C65638135 CRC64;

arity 100.0%; Score 7; Pred. No. 43;
onservative 0; Mismatches 0; Indels 0; Gaps 0;

RCL 227
|||
RCL 100

RESULT 27
LACB_BUBBU
ID   LACB_BUBBU   STANDARD;   PRT;   180 AA.
AC   P02755; 062822;
DT   21-JUL-1986 (Rel. 01, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Beta-lactoglobulin precursor (Beta-LG).
GN   LGB.
OS   Bubalus bubalis (Domestic water buffalo).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bubalus.
OX   NCBI_TaxID=89462;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Mammary gland;
RX   MEDLINE=99304500; PubMed=10376212;
RA   Das P., Jain S., Nayak S., Apparao K.B.C., Totey S.M., Garg L.C.;
RT   "Molecular cloning and sequence analysis of the cDNA encoding
RT   beta-lactoglobulin in Bubalus bubalis.";
RL   DNA Seq. 10:105-108(1999).
RN   [2]
RP   SEQUENCE OF 19-180.
RA   Kolde H.-J., Liberatori J., Braunitzer G.;
RT   "The amino acid sequence of the water buffalo beta-lactoglobulin
RL   Milchwissenschaft 36:83-86(1981).
CC   -!- FUNCTION: Primary component of whey, it binds retinol and is
CC   probably involved in the transport of that molecule.
CC   -!- SUBUNIT: Under physiological conditions beta-lactoglobulin e
CC   as an equilibrium mixture of monomeric and dimeric forms.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Synthesized in mammary glound and secrete
CC   in milk.
CC   -!- PTM: Alternate disulfide bonds occur in equal amounts.
CC   -!- SIMILARITY: Belongs to the lipocalin family.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a col
CC   between the Swiss Institute of Bioinformatics and the EMBL out
CC   the European Bioinformatics Institute. There are no restriction
CC   use by non-profit institutions as long as its content is i
CC   modified and this statement is not removed. Usage by and for
CC   entities requires a license agreement (See http://www.isb-sib.ch
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AJ005429; CAA06532.1; -.
DR   InterPro; IPR002345; Lipocalin.
DR   InterPro; IPR000566; Lipocln_cytFABP.
DR   Pfam; PF00061; lipocalin; 1.
DR   PRINTS; PR00179; LIPOCALIN.
DR   PROSITE; PS00213; LIPOCALIN; 1.
KW   Milk; Whey; Retinol-binding; Transport; Lipocalin; Signal.
FT   SIGNAL   1   18
FT   CHAIN   19   180   BETA-LACTOGLOBULIN.
FT   DISULFID   84   178
FT   DISULFID   124   137
FT   DISULFID   124   139   ALTERNATE.
SQ   SEQUENCE   180 AA; 20023 MW; 6836C97B2C2E33CF CRC64;

Query Match   2.5%; Score 7; DB 1; Length 180;
Best Local Similarity 100.0%; Pred. No. 45;
Matches   7; Conservative 0; Mismatches 0; Indels 0; C

Qy   63 LGLALAC 69
              |||||||
Db   8 LGLALAC 14

RESULT 28
LACB_CAPHI
ID   LACB_CAPHI   STANDARD;   PRT;   180 AA.
AC   P02756;

Best Local Similarity 100.0%; Pred No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; G

Qy 63 LGLALAC 69
|||||||
Db 8 LGLALAC 14

RESULT 29
LACB_SHEEP
ID   LACB_SHEEP   STANDARD;   PRT;   180 AA.
AC   P02757;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Beta-lactoglobulin 1/B, 2/A, and 3/C precursor.
OS   Ovis aries (Sheep), and
OS   Ovis orientalis musimon (Mouflon)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940, 9938;
RN   [1]
RP   SEQUENCE FROM N.A. (BLG 1 AND 2).
RC   SPECIES=Sheep;
RX   MEDLINE=8817249; PubMed=3351935;
RA   Ali S., Clark A.J.;
RT   "Characterization of the gene encoding ovine beta-lactoglobulin.
RT   Similarity to the genes for retinol binding protein and other
RT   secretory proteins.";
RL   J. Mol. Biol. 199:415-426(1988).
RN   [2]
RP   SEQUENCE FROM N.A. (BLG 1).
RC   SPECIES=Sheep;
RX   MEDLINE=87049827; PubMed=3096387;
RA   Gaye P., Hue-Delahaie D., Mercier J.-C., Soulier S., Vilotte J.-L
RA   Furet J.-P.;
RT   "Ovine beta-lactoglobulin messenger RNA: nucleotide sequence and
RT   levels during functional differentiation of the mammary gland.";
RL   Biochimie 68:1097-1107(1986).
RN   [3]
RP   SEQUENCE FROM N.A. (BLG 1).
RC   SPECIES=Sheep;
RX   MEDLINE=89057492; PubMed=3194215;
RA   Harris S., Ali S., Anderson S., Archibald A.L., Clark A.J.;
RT   "Complete nucleotide sequence of the genomic ovine beta-lactoglob
RT   gene.";
RL   Nucleic Acids Res. 16:10379-10380(1988).
RN   [4]
RP   SEQUENCE FROM N.A. (BLG 1 AND 2).
RC   SPECIES=Sheep;
RX   MEDLINE=91007276; PubMed=1976573;
RA   Ali S., McClenaghan M., Simons J.P., Clark A.J.;
RT   "Characterisation of the alleles encoding ovine beta-lactoglobuli
RT   and B.";
RL   Gene 91:201-207(1990).
RN   [5]
RP   SEQUENCE OF 19-180 (BLG 2).
RC   SPECIES=Sheep;
RX   MEDLINE=80219294; PubMed=6155855;
RA   Preaux G., Braunitzer G., Kolde H.-J.;
RT   "Primary structure of ovine beta-lactoglobulin.";
RL   Arch. Int. Physiol. Biochim. 88:B45-B46(1980).
RN   [6]
RP   SEQUENCE OF 19-180 (BLG 3).
RC   SPECIES=Sheep;
RX   MEDLINE=89374823; PubMed=2775495;
RA   Erhardt G., Godovac-Zimmermann J., Conti A.;
RT   "Isolation and complete primary sequence of a new ovine wild-type
RT   beta-lactoglobulin C.";
RL   Biol. Chem. Hoppe-Seyler 370:757-762(1989).
RN   [7]
RP   SEQUENCE OF 19-180 (BLG B).

---

el. 01, Created)
el. 22, Last sequence update)
el. 43, Last annotation update)
bulin precursor (Beta-LG).

(Goat).
azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
heria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
nae; Capra.
25;

N.A.
agagrus; TISSUE=Mammary gland;
559; PubMed=8226387;
oll A., Sanchez A.;
sequencing of the cDNA encoding goat beta-
";
71:2832-2832(1993).

N.A.
, Kim J., Yu M.;
1-1993) to the EMBL/GenBank/DDBJ databases.

N.A.

451; PubMed=7699130;
, A., Sanchez A.;
ence of the caprine beta-lactoglobulin gene.";
77:3493-3497(1994).

9-180.
611; PubMed=511095;
unitzer G., Schrank B., Stangl A.;
d sequence of goat beta-lactoglobulin.";
Z. Physiol. Chem. 360:1595-1604(1979).
Primary component of whey, it binds retinol and is
nvolved in the transport of that molecule.
nder physiological conditions beta-lactoglobulin exists
librium mixture of monomeric and dimeric forms.
R LOCATION: Secreted.
CIFICITY: Synthesized in mammary gland and secreted

nate disulfide bonds occur in equal amounts.
: Belongs to the lipocalin family.

DT entry is copyright. It is produced through a collaboration
Swiss Institute of Bioinformatics and the EMBL outstation -
Bioinformatics Institute. There are no restrictions on its
rofit institutions as long as its content is in no way
this statement is not removed. Usage by and for commercial
res a license agreement (See http://www.isb-sib.ch/announce/
il to license@isb-sib.ch).

CAA41385.1; -.
CAA79623.1; -.
CAA79624.1; -.
CAA83946.1; -.
GGT.
1B00.
002345; Lipocalin.
000566; Lipocln_cytFABP.
lipocalin; 1.
79; LIPOCALIN.
213; LIPOCALIN; 1.
etinol-binding; Transport; Lipocalin; Signal.
1   18
19   180   BETA-LACTOGLOBULIN.
34   178
24   137
24   139   ALTERNATE.
0 AA; 19975 MW; C2449BB02A1A80F1 CRC64;

2.5%; Score 7; DB 1; Length 180;

musimon;
996; PubMed=3426802;
armann J., Conti A., Napolitano L.;
a amino-acid sequence of dimeric beta-lactoglobulin from
ammon musimon) milk.";
oppe-Seyler 368:1313-1319(1987).
: LACTOGLOBULIN IS THE PRIMARY COMPONENT OF WHEY, IT
TINOL AND IS PROBABLY INVOLVED IN THE TRANSPORT OF
ECULE.
Under physiological conditions beta-lactoglobulin exists
ilibrium mixture of monomeric and dimeric forms.
ernate disulfide bonds occur in equal amounts.
TY: Belongs to the lipocalin family.

ROT entry is copyright. It is produced through a collaboration
Swiss Institute of Bioinformatics and the EMBL outstation -
Bioinformatics Institute. There are no restrictions on its
-profit institutions as long as its content is in no way
this statement is not removed. Usage by and for commercial
ires a license agreement (See http://www.isb-sib.ch/announce/
nail to license@isb-sib.ch).

; CAA28204.1; -.
; CAA31305.1; -.
; CAA30059.1; ALT_SEQ.
; CAA30059.1; JOINED.
; CAA30059.1; JOINED.
; CAA30059.1; JOINED.
; CAA30059.1; JOINED.
; CAA30059.1; JOINED.
; AAA31510.1; -.
; AAA31510.1; JOINED.
; AAA31510.1; JOINED.
; AAA31510.1; JOINED.
LGSH.
; 1BSQ.
R002345; Lipocalin.
R000566; LipocIn_cytFABP.
1; lipocalin; 1.
179; LIPOCALIN.
0213; LIPOCALIN; 1.
Retinol-binding; Transport; Signal; Lipocalin.
    19   18   BETA-LACTOGLOBULIN.
    84   178
   124   137   ALTERNATE.
   124   139   H -> Y (IN LACTOGLOBULIN 2=A AND 3=C).
    38   38   R -> Q (IN LACTOGLOBULIN 3=C).
   166   166   BABC2E89E757333 CRC64;
30 AA; 19921 MW;

larity   2.5%;   Score 7; DB 1;   Length 180;
Conservative   0; Mismatches   0; Indels   0; Gaps   0;

ALAC 69
|||||
ALAC 14

         STANDARD;   PRT;   181 AA.

(Rel. 42, Created)
(Rel. 42, Last sequence update)
(Rel. 42, Last annotation update)
phoribosyltransferase (EC 2.4.2.7) (APRT).
2.

neidensis.
oteobacteria; Gammaproteobacteria; Alteromonadales;
ceae; Shewanella.
0863;

RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MR-1;
RX   MEDLINE=2297686; PubMed=12368813;
RA   Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson
RA   Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R
RA   Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty
RA   DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA   Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA   Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA   Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA   Feldblyum T.V., Smith H.O., Venter J.C., Nealson K.H., Fraser C.
RT   "Genome sequence of the dissimilatory metal ion-reducing bacteri
RT   Shewanella oneidensis.";
RL   Nat. Biotechnol. 20:1118-1123(2002).
-!- FUNCTION: Catalyzes a salvage reaction resulting in the form
    of AMP, that is energically less costly than de novo synthes
-!- CATALYTIC ACTIVITY: AMP + diphosphate = adenine + 5-phospho-
    D-ribose 1-diphosphate.
-!- PATHWAY: Purine salvage.
-!- SUBUNIT: Homodimer (By similarity).
-!- SUBCELLULAR LOCATION: Cytoplasmic.
-!- SIMILARITY: Belongs to the purine/pyrimidine
    phosphoribosyltransferase family.

DR   EMBL; AE015643; AAN55062.1; -.
DR   TIGR; SO2012; -.
DR   HAMAP; MF_00004; -; 1.
DR   InterPro; IPR005764; Ade_phspho_trans.
DR   InterPro; IPR002375; Pr/py_rp_transf.
DR   InterPro; IPR000836; PRTransferase.
DR   Pfam; PF00156; Pribosyltran; 1.
DR   TIGRFAMS; TIGR01090; apt; 1.
DR   PROSITE; PS00103; PUR_PYR_PR_TRANSFER; 1.
KW   Transferase; Glycosyltransferase; Purine salvage; Complete prote
SQ   SEQUENCE   181 AA; 19543 MW; C4255A59C4632CA4 CRC64;

Query Match   2.5%;   Score 7; DB 1;   Length 181;
Best Local Similarity 100.0%;   Pred. No. 45;
Matches   7; Conservative   0; Mismatches   0; Indels   0;

Qy   58 PLAIGLG 64
           |||||||
Db   71 PLAIGLG 77

RESULT 31
DEF1_BIFLO
ID   DEF1_BIFLO   STANDARD;   PRT;   217 AA.
AC   Q8G534;
DT   15-MAR-2004 (Rel. 43, Created)
DT   15-MAR-2004 (Rel. 43, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Peptide deformylase 1 (EC 3.5.1.88) (PDF 1) (Polypeptide deformy
DE   1).
GN   DEF1 OR BL1186.
OS   Bifidobacterium longum.
OC   Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
OC   Bifidobacteriaceae; Bifidobacterium.
OX   NCBI_TaxID=216816;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCC 2705;
RX   MEDLINE=2294977; PubMed=12381787;
RA   Schell M.A., Karmirantzou M., Snel B., Vilanova D., Berger B.,

RC   SPECIES=S.typhi; STRAIN=CT18;
RX   MEDLINE=21534947; PubMed=11677608;
RA   Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wai
RA   Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia
RA   Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA   Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA   Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K
RA   Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.
RA   Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA   Whitehead S., Barrell B.G.;
RT   "Complete genome sequence of a multiple drug resistant Salmonella
RT   enterica serovar Typhi CT18.";
RL   Nature 413:848-852(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   SPECIES=S.typhi; STRAIN=Ty2 / ATCC 700931;
RX   MEDLINE=22531367; PubMed=12644504;
RA   Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA   Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT   "Comparative genomics of Salmonella enterica serovar Typhi strain
RT   and CT18.";
RL   J. Bacteriol. 185:2330-2337(2003).
CC   -!- FUNCTION: May be part of a membrane complex involved in elect
CC   transport (By similarity).
CC   -!- SUBUNIT: Composed of at least six subunits; rnfA, rnfB, rnfC,
CC   rnfD, rnfE and rnfG (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membra
CC   (Potential).
CC   -!- SIMILARITY: Belongs to the nqrDE/rnfAB family.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a coll
CC   between the Swiss Institute of Bioinformatics and the EMBL out
CC   the European Bioinformatics Institute. There are no restriction
CC   use by non-profit institutions as long as its content is in
CC   modified and this statement is not removed. Usage by and for c
CC   entities requires a license agreement (See http://www.isb-sib.ch/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AE008763; AAL20376.1; -.
DR   EMBL; AL627271; CAD01913.1; -.
DR   EMBL; AE016838; AAO68972.1; -.
DR   StyGene; SG?????; rnfE.
DR   HAMAP; MF_00478; -; 1.
DR   InterPro; IPR003667; Rnf_Nqr.
DR   Pfam; PF02508; Rnf-Nqr; 1.
KW   Electron transport; Transmembrane; Inner membrane; Complete prote
FT   TRANSMEM   34   56   POTENTIAL.
FT   TRANSMEM   69   87   POTENTIAL.
FT   TRANSMEM   91   113   POTENTIAL.
FT   TRANSMEM   126   148   POTENTIAL.
FT   TRANSMEM   183   205   POTENTIAL.
SQ   SEQUENCE   230 AA; 24318 MW; E198B4CEA13F249E CRC64;

Query Match   2.5%; Score 7; DB 1; Length 230;
Best Local Similarity   100.0%; Pred. No. 56;
Matches   7; Conservative   0; Mismatches   0; Indels   0; C

Qy   60 ALGLGLA 66
               |||||||
Db   38 ALGLGLA 44

RESULT 33
RNFE_ECO57
ID   RNFE_ECO57   STANDARD;   PRT;   231 AA.
AC   P58344;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Electron transport complex protein rnfE.
GN   RNFE OR Z2642 OR ECS2341
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

RA   ...len M.-C., Desiere F., Bork P., Delley M.,
RA   ...Arigoni F.;
RT   ...equence of Bifidobacterium longum reflects its adaptation
RT   gastrointestinal tract.";
RL   ...cad. Sci. U.S.A. 99:14422-14427(2002).
CC   Removes the formyl group from the N-terminal Met of
CC   ...thesized proteins. Requires at least a dipeptide for an
CC   rate of reaction. N-terminal L-methionine is a
CC   ...ite for activity but the enzyme has broad specificity at
CC   ...itions (By similarity).
CC   -!- ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
CC   peptide.
CC   -!- ...Binds 1 iron(II) ion (By similarity).
CC   ...Y: Belongs to the polypeptide deformylase family.
CC   --
CC   ...OT entry is copyright. It is produced through a collaboration
CC   ...Swiss Institute of Bioinformatics and the EMBL outstation -
CC   Bioinformatics Institute. There are no restrictions on its
CC   ...profit institutions as long as its content is in no way
CC   ...this statement is not removed. Usage by and for commercial
CC   ...ires a license agreement (See http://www.isb-sib.ch/announce/
CC   ...ail to license@isb-sib.ch).
CC   --
DR   ...3; AAN4991.1; -.
DR   ...63; -; 1.
DR   ...000181; Fmet deformylase.
DR   ; Pep deformylase; 1.
DR   ...76; PDEFORMYLASE.
KW   ...844; Pep deformylase; 1.
KW   ...nthesis; Hydrolase; Iron; Complete proteome.
FT   72   172   BY SIMILARITY.
FT   29   129   IRON (BY SIMILARITY).
FT   71   171   IRON (BY SIMILARITY).
FT   75   175   IRON (BY SIMILARITY).
SQ   ...7 AA; 24443 MW; 52F5B46B6F47163 CRC64;

...arity   2.5%; Score 7; DB 1; Length 217;
... 100.0%; Pred. No. 53;
...onservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   ALA 68
            |||
Db   ALA 86

ID   ...   STANDARD;   PRT;   230 AA.
DT   Rel. 41, Created)
DT   Rel. 41, Last sequence update)
DT   Rel. 42, Last annotation update)
DE   ...sport complex protein rnfE.
GN   ...54 OR STY1668 OR T1322.
OS   ...phimurium, and
OS   phi.
OC   ...teobacteria; Gammaproteobacteria; Enterobacteriales;
OC   ...aceae; Salmonella.
OC   ...2, 601;
RP   N.A.
RA   ...himurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX   ...948; PubMed=11677609;
RA   , Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA   Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA   ...guyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Wilson R.K.;
RT   ...ome sequence of Salmonella enterica serovar Typhimurium
RL   2-856(2001).
RN   N.A.

```
ID   RNFE_ECOLI     STANDARD;      PRT;    231 AA.
AC   P77179;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Electron transport complex protein rnfE.
GN   RNFE OR B1632.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97251357; PubMed=9097039;
RA   Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA   Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
RA   Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,
RA   Motomura K., Nakade S., Nakamura Y., Nashimoto H., Nishio Y.,
RA   Oshima T., Saito N., Sampei G., Seki Y., Sivasundaram S.,
RA   Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
RA   Yamamoto Y., Horiuchi T.;
RT   "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 28.0-40.1 min region on the linkage map.";
RL   DNA Res. 3:363-377(1996).
RN   [3]
RP   TOPOLOGY.
RC   STRAIN=K12 / JM109.
RX   MEDLINE=99342054; PubMed=10411911;
RA   Saeaef A., Johansson M., Wallin E., von Heijne G.;
RT   "Divergent evolution of membrane protein topology:  the Escherich
RT   coli RnfA and RnfE homologues.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:8540-8544(1999).
CC   -!- FUNCTION: May be part of a membrane complex involved in elec
CC       transport (By similarity).
CC   -!- SUBUNIT: Composed of at least six subunits; rnfA, rnfB, rnfC
CC       rnfD, rnfE and rnfG (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membr
CC       (Potential).
CC   -!- SIMILARITY: Belongs to the nqrDE/rnfAE family.
CC   -------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a col
CC   between the Swiss Institute of Bioinformatics and the EMBL ou
CC   the European Bioinformatics Institute.  There are no restrictio
CC   use by non-profit institutions as long  as its content is i
CC   modified and this statement is not removed.  Usage by and for
CC   entities requires a license agreement (See http://www.isb-sib.ch).
CC   or send an email to license@isb-sib.ch).
CC   -------------------------------------------------------------------
DR   EMBL; AE000258; AAC74704.1; -.
DR   EMBL; D90806; BAA15386.1; -.
DR   EMBL; D90807; BAA15393.1; -.
DR   EMBL; D90808; BAA15416.1; -.
DR   PIR; B64920; B64920.
DR   EcoGene; EG13938; rnfE.
DR   HAMAP; MF_00478; -; 1.
DR   InterPro; IPR003667; Rnf_Nqr.
DR   Pfam; PF02508; Rnf-Nqr; 1.
KW   Electron transport; Transmembrane; Inner membrane; Complete prot
FT   DOMAIN        1     38       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM     39     59       POTENTIAL.
FT   DOMAIN       60     62       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     63     83       POTENTIAL.
```

```
aceae; Escherichia.
334;
N.A.
H7 / EDL933 / ATCC 700927;
935; PubMed=11206551;
Plunkett G. III, Burland V., Mau B., Glasner J.D.,
ayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
ckett J., Klink S., Boutin A., Shao Y., Miller L.,
., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
Blattner F.R.;
ence of enterohaemorrhagic Escherichia coli O157:H7.";
29-533(2001).
N.A.
H7 / RIMD 0509952;
5231; PubMed=11258796;
Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
itsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
ami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
iiba T., Hattori M., Shinagawa H.;
ome sequence of enterohemorrhagic Escherichia coli
genomic comparison with a laboratory strain K-12.";
-22(2001).
TY: Belongs to the nqrDE/rnfAE family.
--------------------------------------------------------------
ROT entry is copyright. It is produced through a collaboration
Swiss Institute of Bioinformatics and the EMBL outstation -
Bioinformatics Institute.  There are no restrictions on its
-profit institutions as long as its content is in no way
this statement is not removed.  Usage by and for commercial
uires a license agreement (See http://www.isb-sib.ch/announce/
nail to license@isb-sib.ch).
--------------------------------------------------------------
36; AAG56621.1; -.
58; BAB35764.1; -.
A85770.
E90921.
178; -; 1.
R003667; Rnf_Nqr.
3; Rnf-Nqr; 1.
nsport; Transmembrane; Inner membrane; Complete proteome.
1     38       PERIPLASMIC (POTENTIAL).
39     59       POTENTIAL.
60     62       CYTOPLASMIC (POTENTIAL).
63     83       PERIPLASMIC (POTENTIAL).
86    106       POTENTIAL.
107    124       CYTOPLASMIC (POTENTIAL).
125    145       POTENTIAL.
146    181       PERIPLASMIC (POTENTIAL).
182    202       POTENTIAL.
203    231       CYTOPLASMIC (POTENTIAL).
31 AA;  24489 MW;  D4A2CA2D292604C3 CRC64;
```

```
larity:    2.5%;   Score 7;   DB 1;  Length 231;
                                      Pred. No. 56;
Conservative    0;  Mismatches    0;   Indels    0;  Gaps    0;

LGLA 66
|||
LGLA 44
```

```
DR   Pfam; PF00035; dsrm; 1.
DR   Pfam; PF00636; Ribonuclease_3; 1.
DR   SMART; SM00358; DSRM; 1.
DR   SMART; SM00535; RIBOc; 1.
DR   PROSITE; PS50137; DS_RBD; 1.
DR   PROSITE; PS00517; RNASE_3_1; 1.
DR   PROSITE; PS50142; RNASE_3_2; 1.
KW   Hydrolase; Nuclease; Endonuclease; RNA-binding; Complete proteome
FT   DOMAIN        4    126       RNASE III.
FT   DOMAIN      204    220       DRBM.
FT   CONFLICT    116    116       A -> T (IN REF. 1).
SQ   SEQUENCE   233 AA;  26199 MW;  1A11CB5FD960784F CRC64.
```

```
Query Match             2.5%;  Score 7;  DB 1;  Length 233;
Best Local Similarity 100.0%;  Pred. No. 56;
Matches    7;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;
```

```
Qy      10 ARRLPLP 16
           |||||||
Db     164 ARRLPLP 170
```

```
RESULT 36
RNFE_YERPE
ID   RNFE_YERPE     STANDARD;     PRT;     233 AA.
AC   Q8ZED4;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Electron transport complex protein rnfE.
GN   RNFE OR YPO2240 OR Y2081.
OS   Yersinia pestis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Yersinia.
OX   NCBI_TaxID=632;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CO-92 / Biovar Orientalis;
RX   MEDLINE=21470413; PubMed=11586360;
RA   Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA   Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,
RA   Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA   Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA   Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA   Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
RA   Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT   "Genome sequence of Yersinia pestis, the causative agent of plague
RL   Nature 413:523-527(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KIM5 / Biovar Mediaevalis;
RX   MEDLINE=22137863; PubMed=12142430;
RA   Deng W., Burland V., Plunkett G. III, Boutin A., M
RA   Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz [
RA   Fetherston J.D., Lindler L.E., Brubaker R.R., Pla
RA   Straley S.C., McDonough K.A., Nilles M.L., Matso
RA   Perry R.D.;
RT   "Genome sequence of Yersinia pestis KIM.";
RL   J. Bacteriol. 184:4601-4611(2002).
CC   -!- FUNCTION: May be part of a membrane comple
CC       transport (By similarity).
CC   -!- SUBUNIT: Composed of at least six subunit
CC       rnfD, rnfE and rmfG (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane 
CC       (Potential).
CC   -!- SIMILARITY: Belongs to the nqrDE/rnfAE 
CC   --------------------------------------------
CC   This SWISS-PROT entry is copyright. It is [
CC   between the Swiss Institute of Bioinforma
CC   the European Bioinformatics Institute. Th
CC   use by non-profit institutions as long
CC   modified and this statement is not remov
CC   entities requires a license agreement (S
```

Left column (rotated):

```
4     85       PERIPLASMIC (POTENTIAL).
6    106       POTENTIAL.
7    124       CYTOPLASMIC (POTENTIAL).
5    145       POTENTIAL.
6    181       PERIPLASMIC (POTENTIAL).
2    202       POTENTIAL.
3    231       CYTOPLASMIC (POTENTIAL).
. AA;  24459 MW;  CFA37A2D292604C3 CRC64;
```

```
urity       2.5%;  Score 7;  DB 1;  Length 231;
     100.0%;  Pred. No. 56;
onservative   0;  Mismatches   0;  Indels   0;  Gaps   0;
```

```
:LA 66
    ||
:LA 44
```

```
STANDARD;     PRT;     233 AA.
```

```
Rel. 34, Created)
Rel. 42, Last sequence update)
Rel. 42, Last annotation update)
II (EC 3.1.26.3) (RNase III).
3..
atii.
teobacteria; Gammaproteobacteria; Legionellales;
Coxiella.
7;
```

```
N.A.
7;
751; PubMed=7830573;
ver T.A., Powell B.S., Court D.L.;
the rnc locus of Coxiella burnetii.";
L. 14:291-300(1994).
```

```
N.A.
ile phase I / RSA 493;
557; PubMed=12704232;
Paulsen I.T., Eisen J.A., Read T.D., Nelson K.E.,
ward N.L., Tetelin H., Davidsen T.M., Beanan M.J.,
daugherty S.C., Brinkac L.M., Madupu R., Dodson R.J.,
ee K.H., Carty H.A., Scanlan D., Heinzen R.A.,
Samuel J.E., Fraser C.M., Heidelberg J.F.;
ome sequence of the Q-fever pathogen, Coxiella
```

```
cad. Sci. U.S.A. 100:5455-5460(2003).
Digests double-stranded RNA. Involved in the processing
nal RNA precursors and of some mRNAs (By similarity).
ACTIVITY: Endonucleolytic cleavage to 5'-
noester.
AR LOCATION: Cytoplasmic.
Y: Contains 1 DRBM (double-stranded RNA-binding) domain.
Y: Contains 1 RNase III domain.
--------------------------------------------
OT entry is copyright. It is produced through a collaboration -
Swiss Institute of Bioinformatics and the EMBL outstation -
Bioinformatics Institute. There are no restrictions on its
profit institutions as long as its content is in no way
this statement is not removed. Usage by and for commercial
ires a license agreement (See http://www.isb-sib.ch/announce/
ail to license@isb-sib.ch).
```

```
AAA69690.1; -.
4; AAO91000.1; -.
S60767.
; -; 1.
04; -; 1.
00159; DS_RBD.
000999; RNase_III.
```

CC   nail to license@isb-sib.ch).

DR   HAMAP; MF_00478; -; 1.
DR   InterPro; IPR003667; Rnf_Nqr.
DR   Pfam; PF02508; Rnf-Nqr; 1.
KW   Electron transport; Transmembrane; Inner membrane; Complete prot
FT   TRANSMEM   63   83   POTENTIAL.
FT   TRANSMEM   93   113   POTENTIAL.
FT   TRANSMEM   117   137   POTENTIAL.
FT   TRANSMEM   152   172   POTENTIAL.
FT   TRANSMEM   206   226   POTENTIAL.
SQ   SEQUENCE   235 AA;   25845 MW; CO54FE59647837A CRC64;

Query Match   2.5%; Score 7; DB 1; Length 235;
Best Local Similarity 100.0%;   Pred. No. 57;
Matches   7; Conservative   0; Mismatches   0; Indels   0;

Qy   60 ALGLGLA 66
             |||||||
Db   62 ALGLGLA 68

RESULT 38
ID   TN14_MOUSE   STANDARD;   PRT;   239 AA.
AC   Q9QYH9;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Tumor necrosis factor ligand superfamily member 14.
GN   TNFSF14 OR LIGHT.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mu
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20165223; PubMed=10700230;
RA   Tamada K., Shimozaki K., Chapoval A.I., Zhu G., Sica G., Flies D
RA   Boone T., Hsu H., Fu Y.-X., Nagata S., Ni J., Chen L.;
RT   "Modulation of T-cell-mediated immunity in tumor and graft-versu
RT   disease models through the LIGHT co-stimulatory pathway.";
RL   Nat. Med. 6:283-289(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal liver;
RX   MEDLINE=20354998; PubMed=10894944;
RA   Misawa K., Nosaka T., Kojima T., Hirai M., Kitamura T.;
RT   "Molecular cloning and characterization of a mouse homolog of hu
RT   TNFSF14, a member of the TNF superfamily.";
RL   Cytogenet. Cell Genet. 89:89-91(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lymphoma;
RA   Force W.R., Todd P.K., Mikayama T.;
RT   "Mouse LIGHT, molecular genetics, ligand binding and expression.
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Cytokine that binds to TNFRSF3/LTBR. Binding to th
CC   decoy receptor TNFRSF6B modulates its effects. Activates NFK
CC   and stimulates the proliferation of T cells.
CC   -!- SUBUNIT: Homotrimer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein and secreted
CC   similarity).
CC   -!- PTM: The soluble form derives from the membrane form by
CC   proteolytic processing.
CC   -!- SIMILARITY: Belongs to the tumor necrosis factor family.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a col
CC   between the Swiss Institute of Bioinformatics and the EMBL ou
CC   the European Bioinformatics Institute. There are no restrictio
CC   use by non-profit institutions as long as its content is
CC   modified and this statement is not removed. Usage by and for
CC   entities requires a license agreement (See http://www.isb-sib.ch
CC   or send an email to license@isb-sib.ch).
CC   --

CC   nail to license@isb-sib.ch).

DR   AB0273.
DR   51; CAC91046.1; -.
DR   10; AAM85645.1; -.
DR   478; -; 1.
DR   R003667; Rnf_Nqr.
DR   8; Rnf-Nqr; 1.
KW   nsport; Transmembrane; Inner membrane; Complete proteome.
FT   TRANSMEM   34   56   POTENTIAL.
FT   TRANSMEM   69   87   POTENTIAL.
FT   TRANSMEM   91   113   POTENTIAL.
FT   TRANSMEM   126   148   POTENTIAL.
FT   TRANSMEM   184   206   POTENTIAL.
SQ   SEQUENCE   233 AA;   24587 MW; 491E18F335E8CB90 CRC64;

Query Match   2.5%; Score 7; DB 1; Length 233;
Best Local Similarity 100.0%;   Pred. No. 56;
Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   66 LGLA 66
             ||||
Db   44 LGLA 44

ID   STANDARD;   PRT;   235 AA.
AC   46;
DT   (Rel. 35, Created)
DT   (Rel. 35, Last sequence update)
DT   (Rel. 41, Last annotation update)
DE   nsport complex protein rnfE.
OS   88.
OC   influenzae.
OC   oteobacteria; Gammaproteobacteria; Pasteurellales;
OC   eae; Haemophilus.
OX   27;
RN
RP   N.A.
RC   M.N.A.
RC   KW20 / ATCC 51907;
RX   0630; PubMed=7542800;
RA   R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA   R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA   Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA   Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA   , Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA   R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA   ritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA   McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RT   e random sequencing and assembly of Haemophilus influenzae
RL   496-512(1995).
CC   -!- May be part of a membrane complex involved in electron
CC   t. (By similarity).
CC   -!- Composed of at least six subunits; rnfA, rnfB, rnfC,
CC   fE and rnfG (By similarity).
CC   LAR LOCATION: Integral membrane protein. Inner membrane
CC   al).
CC   TY: Belongs to the nqrDE/rnfAE family.
CC   --
CC   ROT entry is copyright. It is produced through a collaboration
CC   Swiss Institute of Bioinformatics and the EMBL outstation -
CC   Bioinformatics Institute. There are no restrictions on its
CC   -profit institutions as long as its content is in no way
CC   this statement is not removed. Usage by and for commercial
CC   uires a license agreement (See http://www.isb-sib.ch/announce/
CC   mail to license@isb-sib.ch).
CC   --
DR   , AAC23334.1; -.
DR   I64174.
DR   ; -.

CC   This SWISS-PROT entry is copyright. It is produced through a coll
CC   between the Swiss Institute of Bioinformatics and the EMBL out
CC   the European Bioinformatics Institute. There are no restriction
CC   use by non-profit institutions as long as its content is in
CC   modified and this statement is not removed. Usage by and for c
CC   entities requires a license agreement (See http://www.isb-sib.ch/
CC   or send an email to license@isb-sib.ch).
DR   EMBL; AE004770; AAG06882.1; -.
DR   PIR; G83208; G83208.
DR   HAMAP; MF_00478; -; 1.
DR   InterPro; IPR003667; Rnf_Nqr.
DR   Pfam; PF02508; Rnf-Nqr; 1.
KW   Electron transport; Transmembrane; Inner membrane; Complete prote
FT   TRANSMEM   41   61   POTENTIAL.
FT   TRANSMEM   71   91   POTENTIAL.
FT   TRANSMEM   95   115   POTENTIAL.
FT   TRANSMEM   130   150   POTENTIAL.
FT   TRANSMEM   184   204   POTENTIAL.
SQ   SEQUENCE   240 AA; 25794 MW; 3D9O6B7ED462DBB2 CRC64;

Query Match   2.5%; Score 7; DB 1; Length 240;
Best Local Similarity   100.0%; Pred. No. 58;
Matches   7; Conservative   0; Mismatches   0; Indels   0; G

Qy   60 ALGLGLA 66
            |||||||
Db   40 ALGLGLA 46

RESULT 40
MOEB_HAEIN
ID   MOEB_HAEIN   STANDARD;   PRT;   243 AA.
AC   P45211;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Molybdopterin biosynthesis protein moeB.
GN   MOEB OR CHIN OR HI1449.
OS   Haemophilus influenzae.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Haemophilus.
OX   NCBI_TaxID=727;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Rd / KW20 / ATCC 51907;
RX   MEDLINE=95350630; PubMed=7542800;
RA   Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.
RA   Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.
RA   McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA   Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA   Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA   Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C
RA   Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA   Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA   Venter J.C.;
RT   "Whole-genome random sequencing and assembly of Haemophilus influ
RT   Rd.";
RL   Science 269:496-512(1995).
CC   -!- FUNCTION: INVOLVED IN BIOSYNTHESIS OF A DEMOLYBDO COFACTOR
CC   (MOLYBDOPTERIN), NECESSARY FOR MOLYBDOENZYMES. PLAYS A ROLE I
CC   ACTIVATION OF THE SMALL SUBUNIT OF THE MOLYBDOPTERIN CONVERTI
CC   FACTOR (MOAD) (BY SIMILARITY).
CC   -!- PATHWAY: Molybdenum cofactor biosynthesis.
CC   -!- SIMILARITY: BELONGS TO THE HESA/MOEB/THIF FAMILY.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a coll
CC   between the Swiss Institute of Bioinformatics and the EMBL out
CC   the European Bioinformatics Institute. There are no restriction
CC   use by non-profit institutions as long as its content is in
CC   modified and this statement is not removed. Usage by and for c
CC   entities requires a license agreement (See http://www.isb-sib.ch).
CC   or send an email to license@isb-sib.ch).

5; AAF76453.1; -.
5; BAA88559.1; -.
3; AAF36722.1; -.
4TSV.
317; Tnfsf14.
006053; TNF_abc.
006052; TNF family.
008983; TNF_like.
003636; TNF_subf.
; TNF; 1.
34; TNECROSISFCT
012; TNF_subf; 1.
7; TNF; 1.
251; TNF_1; FALSE_NEG.
049; TNF_2; 1.
asmembrane; Glycoprotein; Signal-anchor.
1   239   TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
                     MEMBER 14, MEMBRANE FORM.
82   239   TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
                     MEMBER 14, SOLUBLE FORM.
1   37   CYTOPLASMIC (POTENTIAL).
38   58   SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
                     (POTENTIAL).
59   239   EXTRACELLULAR (POTENTIAL).
31   82   CLEAVAGE (POTENTIAL).
52   187   POTENTIAL.
00   100   N-LINKED (GLCNAC. . .) (POTENTIAL).
91   191   N-LINKED (GLCNAC. . .) (POTENTIAL).
9 AA;   26338 MW; 217874AC71AD6BE3 CRC64;

arity   2.5%; Score 7; DB 1; Length 239;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

RPR 127
|||
RPR 227

STANDARD;   PRT;   240 AA.

Rel. 41, Created)
Rel. 41, Last sequence update)
Rel. 41, Last annotation update)
sport complex protein rnfE.
4.
eruginosa.
ceobacteria; Gammaproteobacteria; Pseudomonadales;
eae; Pseudomonas.
7;
N.A.
5692 / PAO1;
337; PubMed=10984043;
Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
oulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
pencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
ier M.H., Hancock R.E.W., Lory S., Olson M.V.;
ome of Pseudomonas aeruginosa PAO1, an
pathogen.";
9-964(2000).
May be part of a membrane complex involved in electron
(By similarity).
Composed of at least six subunits; rnfA, rnfB, rnfC,
E and rnfG (By similarity).
AR LOCATION: Integral membrane protein. Inner membrane
l). Belongs to the ngrDE/rnfAE family.
Y: Belongs to the ngrDE/rnfAE family.

; AAC23099.1; -.
C64124.
-...-
R009036; MoeB.
R007901; MoeZ_MoeB.
R000205; NAD_BS.
R000594; ThiF_domain.
7; MoeZ_MoeB; 1.
9; ThiF; 1.
ofactor biosynthesis; Complete proteome.
13 AA; 26996 MW; 218A3382A975BDBD CRC64;

larity   2.5%;   Score 7;   DB 1;   Length 243;
          100.0%; Pred No.59;
onservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

NRQI 185
||||
NRQI 16

STANDARD;   PRT;   244 AA.
70; Q99761;
(Rel. 29, Created)
(Rel. 29, Last sequence update)
(Rel. 43, Last annotation update)
beta (LT-beta) (Tumor necrosis factor C) (TNF-C) (Tumor
tor ligand superfamily member 3).
3 OR TNFC.
(Human).
etazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
cheria; Primates; Catarrhini; Hominidae; Homo.
506;

M N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.
1;
8881; PubMed=7916655;
Ngam-Ek A., Lawton P., Demarinis J., Tizard R.,
ession C., O'Brine-Greco B., Foley S.F., Ware C.F.;
beta, a novel member of the TNF family that forms a
complex with lymphotoxin on the cell surface.";
356(1993).

M N.A. (ISOFORMS 1 AND 2).
5965; PubMed=9299492;
Renard N., Charlot C., Bienvenu J., Coiffier B.,

ion of two lymphotoxin beta isoforms expressed in human
1 lines and non-Hodgkin's lymphomas.";
phys. Res. Commun. 238:273-276(1997).

M N.A. (ISOFORM 1).
, Milner C.M., Campbell R.D.;
r of the immunoglobulin superfamily and a V-ATPase G
amongst the predicted products of novel genes close to the
the human MHC.";
EP-1997) to the EMBL/GenBank/DDBJ databases.

M N.A. (ISOFORM 1).
dan A., Qin S., Shaffer T., James R., Ratcliffe A.,
ickhoff R., Loretz C., Madan A., Dors M., Young J.,
od L.;
the human major histocompatibility complex class III
CT-1999) to the EMBL/GenBank/DDBJ databases.

M N.A. (ISOFORM 1).
amiya G., Oka A., Inoko H.;
s 2,229,817bp genomic DNA of 6p21.3 HLA class I region.";
EP-1999) to the EMBL/GenBank/DDBJ databases.

RN   [6]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANTS GLU-70 AND
RP   PRO-111.
RA   Rieder M.J., Armel T.Z., Carrington D.P., Chung M.-W., Lee K.L.,
RA   Poel C.L., Toth E.J., Yi Q., Nickerson D.A.;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANTS ARG-84 AND PHE-87.
RA   Rieder M.J., Livingston R.J., Daniels M.R., Montoya M.A., Chung
RA   Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.
RA   Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Cytokine that binds to LTBR/TNFRSF3. May play a sp
CC   role in immune response regulation. Provides the membrane an
CC   for the attachment of the heterotrimeric complex to the cell
CC   surface. Isoform 2 is probably non-functional.
CC   -!- SUBUNIT: Heterotrimer of either two LTB and one LTA subunits
CC   (less prevalent) one LTB and two LTA subunits.
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
CC   -!- ALTERNATIVE PRODUCTS:
CC   Event=Alternative splicing; Named isoforms=2;
CC   Name=1;
CC   IsoId=Q06643-1; Sequence=Displayed;
CC   Name=2;
CC   IsoId=Q06643-2; Sequence=VSP_006441, VSP_006442;
CC   -!- TISSUE SPECIFICITY: Spleen and thymus.
CC   -!- SIMILARITY: Belongs to the tumor necrosis factor family.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a col
CC   between the Swiss Institute of Bioinformatics and the EMBL ou
CC   the European Bioinformatics Institute. There are no restrictio
CC   use by non-profit institutions as long as its content is i
CC   modified and this statement is not removed. Usage by and for
CC   entities requires a license agreement (See http://www.isb-sib.ch
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; L11016; AAA99888.1; -.
DR   EMBL; U89922; AAC51769.1; -.
DR   EMBL; U79029; AAB37342.1; -.
DR   EMBL; L11015; AAA36191.1; -.
DR   EMBL; Y14768; CAA75069.1; -.
DR   EMBL; AF129756; AAD18089.1; -.
DR   EMBL; AP000505; BAB63395.1; -.
DR   EMBL; AY070219; AAL49954.1; -.
DR   EMBL; AY070219; AAL49955.1; -.
DR   EMBL; AY216497; AAO21134.1; -.
DR   PIR; A46066; A46066.
DR   PIR; JC5645; JC5645.
DR   HSSP; P01374; 1TNR.
DR   Genew; HGNC:6711; LTB.
DR   MIM; 600978;
DR   GO; GO:0005102; F:receptor binding; TAS.
DR   GO; GO:0015070; F:toxin activity; NAS.
DR   GO; GO:0007267; P:cell-cell signaling; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR006053; TNF abc.
DR   InterPro; IPR006052; TNF_family.
DR   InterPro; IPR008983; TNF_like.
DR   InterPro; IPR003636; TNF_subf.
DR   Pfam; PF00229; TNF; 1.
DR   PRINTS; PR01234; TNECROSISFCT.
DR   ProDom; PD002012; TNF_subf; 1.
DR   SMART; SM00207; TNF; 1.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 1.
KW   Cytokine; Transmembrane; Glycoprotein; Signal-anchor;
KW   Alternative splicing; Polymorphism.
FT   DOMAIN   1   18   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   19   48   SIGNAL-ANCHOR (TYPE-II MEMBRANE PRO
FT   (POTENTIAL).
FT   DOMAIN   49   244   EXTRACELLULAR (POTENTIAL).
FT   CARBOHYD   222   222   N-LINKED (GLCNAC...) (POTENTIAL).
FT   VARSPLIC   53   77   GLVTETADPGAQQGLGFQKLPEE -> GLGFRS

CC   ---
DR   EMBL; AB054536; BAB83881.1; -.
DR   EMBL; AB100082; BAC78156.1; -.
DR   InterPro; IPR006053; TNF_abc.
DR   InterPro; IPR006052; TNF_family.
DR   InterPro; IPR008983; TNF_like.
DR   InterPro; IPR003636; TNF_subf.
DR   Pfam; PF00229; TNF; 1.
DR   PRINTS; PR01234; TNECROSISFCT.
DR   ProDom; PD002012; TNF_subf; 1.
DR   SMART; SM00207; TNF; 1.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 1.
KW   Cytokine; Transmembrane; Glycoprotein; Signal-anchor.
FT   DOMAIN   1   18   SIGNAL-ANCHOR (TYPE-II MEMBRANE PROT
FT   (POTENTIAL).
FT   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   19   48   SIGNAL-ANCHOR (TYPE-II MEMBRANE PROT
FT   (POTENTIAL).
FT   DOMAIN   49   244   EXTRACELLULAR (POTENTIAL).
FT   CARBOHYD   222   222   N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   244 AA; 25420 MW; A04785833505B97 CRC64;

Query Match   2.5%; Score 7; DB 1; Length 244;
Best Local Similarity 100.0%; Pred. No. 59;
Matches   7; Conservative   0; Mismatches   0; Indels   0; G

Qy 194 GLYYLYC 200
        |||||||
Db 131 GLYYLYC 137

RESULT 43
LFTR_XANCP
ID   LFTR_XANCP   STANDARD;   PRT;   249 AA.
AC   Q8P996;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Leucyl/phenylalany-tRNA--protein transferase (EC 2.3.2.6) (L/F-
DE   transferase) (Leucyltransferase) (Phenylalanyltransferase).
GN   AAT OR XCC1969.
OS   Xanthomonas campestris (pv. campestris).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xanthomonas.
OX   NCBI_TaxID=340;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 33913 / NCPPB 528;
RX   MEDLINE=22022145; PubMed=12024217;
RA   da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R
RA   Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.
RA   Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA   Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.
RA   Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.
RA   Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA   Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA   Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F
RA   Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA   Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA   Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.
RA   Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA   Spinola L.A.F., Takita M.A., Tamura R.B., Teixeira E.C., Tezza R.
RA   Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA   Setubal J.C., Kitajima J.P.;
RT   "Comparison of the genomes of two Xanthomonas pathogens with diff
RT   host specificities.";
RL   Nature 417:459-463(2002).
CC   -!- FUNCTION: Functions in the N-end rule pathway of protein
CC   degradation where it conjugates Leu, Phe and, less efficient
CC   Met from aminoacyl-tRNAs to the N-termini of proteins contain
CC   an N-terminal arginine or lysine (By similarity).
CC   -!- CATALYTIC ACTIVITY: L-leucyl-tRNA + protein = tRNA + L-leucyl
CC   protein.
CC   -!- CATALYTIC ACTIVITY: L-phenylalanyl-tRNA + protein = tRNA + L-

---

KQISAPGSQLPTS (in isoform 2).
/FTId=VSP_006441.
Missing (in isoform 2).
/FTId=VSP_006442.
G -> E.
/FTId=VAR_013025.
S -> R.
/FTId=VAR_016331.
L -> F.
/FTId=VAR_016332.
A -> P.
/FTId=VAR_013026.
DPGAQAQQGL -> GLSAPGSGRT (IN REF. 2;
AAB37342).
AA; 25390 MW; F415694S9830ED4C CRC64;

2.5%; Score 7; DB 1; Length 244;
100.0%; Pred. No. 59;
nservative   0; Mismatches   0; Indels   0; Gaps   0;

YC 200
||
YC 137

STANDARD;   PRT;   244 AA.

Rel. 43, Created)
Rel. 43, Last sequence update)
Rel. 43, Last annotation update)
ta (LT-beta) (Tumor necrosis factor C) (TNF-C) (Tumor
or ligand superfamily member 3).
OR TNFC.
es (Chimpanzee).
azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
teria; Primates; Catarrhini; Hominidae; Pan.
98;

N.A.
002; PubMed=12493009;
hiina T., Anzai T., Kohara S., Inoko H.;
enomic analysis of the MHC: the evolution of class I
ocks, diversity and complexity from shark to man.";
190:95-122(2002).

N.A.
34; PubMed=12799463;
na T., Kimura N., Yanagiya K., Kohara S., Shigenari A.,
ulski J.K., Naruse T.K., Fujimori Y., Fukuzumi Y.,
ashiro H., Iwamoto C., Umehara Y., Imanishi T.,
K., Gojobori T., Bahram S., Inoko H.;
sequencing of human and chimpanzee MHC class I regions
ions/deletions as the major path to genomic

ad. Sci. U.S.A. 100:7708-7713(2003).
Cytokine that binds to LTBR/TNFRSF3. May play a specific
mmune response regulation. Provides the membrane anchor
tachment of the heterotrimeric complex to the cell
By similarity).
heterotrimer of either two LTB and one LTA subunits or
valent) two LTA and one LTB subunits (By similarity).
AR LOCATION: Type II membrane protein (Potential).
Y: Belongs to the tumor necrosis factor family.

CC   (MOLYBDOPTERIN), NECESSARY FOR MOLYBDOENZYMES. PLAYS A ROLE
CC   ACTIVATION OF THE SMALL SUBUNIT OF THE MOLYBDOPTERIN CONVERT
CC   FACTOR (MOAD).
CC   -!- PATHWAY: Molybdenum cofactor biosynthesis.
CC   -!- SIMILARITY: BELONGS TO THE HESA/MOEB/THIF FAMILY.
CC   This SWISS-PROT entry is copyright. It is produced through a col
CC   between the Swiss Institute of Bioinformatics and the EMBL ou
CC   the European Bioinformatics Institute. There are no restrictio
CC   use by non-profit institutions as long as its content is i
CC   modified and this statement is not removed. Usage by and for
CC   entities requires a license agreement (See http://www.isb-sib.ch
CC   or send an email to license@isb-sib.ch).
DR   EMBL; M21151; AAA23580.1; -.
DR   EMBL; AE000185; AAC73913.1; -.
DR   EMBL; D90720; BAA35514.1; -.
DR   EMBL; D90721; BAA35521.1; -.
DR   PIR; B32352; B32352.
DR   PDB; 1JW9; 21-NOV-01.
DR   PDB; 1JWA; 21-NOV-01.
DR   PDB; 1JWB; 21-NOV-01.
DR   EcoGene; EG10154; moeB.
DR   InterPro; IPR009036; MoeB.
DR   InterPro; IPR007901; MoeZ_MoeB.
DR   InterPro; IPR000205; NAD_BS.
DR   InterPro; IPR000594; ThiF_domain.
DR   Pfam; PF05237; MoeZ_MoeB_1.
DR   Pfam; PF00899; ThiF; 1.
KW   Molybdenum cofactor biosynthesis; Complete proteome; 3D-structur
SQ   SEQUENCE   249 AA; 26719 MW; 12C77082B3F39D7D CRC64;

Query Match   2.5%; Score 7; DB 1; Length 249;
Best Local Similarity 100.0%; Pred. No. 60;
Matches   7; Conservative   0; Mismatches   0; Indels   0;

Qy   179 LRYNRQI 185
              |||||||
Db   10 LRYNRQI 16

RESULT 45
MOEB_SALTY
ID   MOEB_SALTY   STANDARD;   PRT;   249 AA.
AC   Q56067;
DT   01-NOV-1997 (Rel. 35, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Molybdopterin biosynthesis protein moeB.
GN   MOEB OR STM0845.
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2;
RA   Wong K.K., Kwan H.S.;
RL   Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX   MEDLINE=21534948; PubMed=11677609;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreill
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Laym
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimu
RT   LT2.";
RL   Nature 413:852-856(2001).
CC   -!- FUNCTION: INVOLVED IN BIOSYNTHESIS OF A DEMOLYBDO COFACTOR
CC   (MOLYBDOPTERIN), NECESSARY FOR MOLYBDOENZYMES. PLAYS A ROLE

---

any-protein.
AR   LOCATION: Cytoplasmic (By similarity).
Y: Belongs to the L/F-transferase family.

OT entry is copyright. It is produced through a collaboration
Swiss Institute of Bioinformatics and the EMBL outstation -
Bioinformatics Institute. There are no restrictions on its
this statement is not removed. Usage by and for commercial
profit institutions as long as its content is in no way
ires a license agreement (See http://www.isb-sib.ch/announce/
ail to license@isb-sib.ch).

01; AAM41258.1; -.
588; -; 1.
R004616; Aat.
3; Leu Phe trans; 1.
R00667; aat; 1.
Acyltransferase; Complete proteome.
19 AA; 27404 MW; E4584D8A164EDB54 CRC64;

larity   2.5%; Score 7; DB 1; Length 249;
Conservative   100.0%; Pred.No. 60;
Conservative   0; Mismatches   0; Indels   0; Gaps   0;

SQDG 156
     |||
SQDG 116

STANDARD;   PRT;   249 AA.

(Rel. 12, Created)
(Rel. 12, Last sequence update)
(Rel. 42, Last annotation update)
n biosynthesis protein moeB.
OR B0826.
coli.
oteobacteria; Gammaproteobacteria; Enterobacteriales;
iaceae; Escherichia.
52;

M N.A.
4906; PubMed=3045084;
sai Y., Saito T.;
sequencing of the Escherichia coli chlEN operon involved
erin biosynthesis.";
. 170:4097-4102(1988).

M N.A.
MG1655;
5617; PubMed=9278503;
.; Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
llado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
avis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Y.;
e genome sequence of Escherichia coli K-12.";
1453-1474(1997).

M N.A.
1202; PubMed=8905232;
iba H., Baba T., Fujita K., Hayashi K., Honjo A.,
Inata H., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
itagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
omura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
eki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
iuchi T.;
A sequence of the Escherichia coli K-12 genome
g to the 12.7-28.0 min region on the linkage map.";
37-155(1996).
: INVOLVED IN BIOSYNTHESIS OF A DEMOLYBDO COFACTOR

OF THE SMALL SUBUNIT OF THE MOLYBDOPTERIN CONVERTING
AD).
olybdenum cofactor biosynthesis.
: BELONGS TO THE HESA/MOEB/THIF FAMILY.

AAA96530.1; -.
5; AAL19781.1; -.
565; moeB.
D09036; MoeB.
D07901; MoeZ MoeB.
D00205; NAD_BS.
D00594; ThiF_domain.
; MoeZ MoeB; 1.
; ThiF_1.
Factor biosynthesis; Complete proteome.
9   9   M -> I (IN REF. 1).
38   38   G -> R (IN REF. 1).
17   117   S -> A (IN REF. 1).
59   169   N -> T (IN REF. 1).
21   221   G -> E (IN REF. 1).
9 AA;   26903 MW; 0F0050B31D537AD2 CRC64;

arity   2.5%;   Score 7; DB 1;   Length 249;
       100.0%; Pred. No. 60;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

RQI 185
RQI 16

STANDARD;   PRT;   257 AA.
Rel. 39, Created)
Rel. 39, Last sequence update)
Rel. 41, Last annotation update)
aride core biosynthesis glycosyl transferase kdtx
escens
eobacteria; Gammaproteobacteria; Enterobacteriales;
aceae; Serratia.
5.
N.A.
003; PubMed=8824620;
Pique N., Climent N., Ferrer S., Merino S., Rubires X.,
ague M.;
characterization of two Serratia marcescens genes
ore lipopolysaccharide biosynthesis.";
178:5741-5747(1996).
Lipopolysaccharide core biosynthesis.
Y: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2. WAAE/KDTX

CC   --
DR   EMBL; U52844; AAC44433.1; -.
DR   InterPro; IPR001173; Glyco_trans 2.
DR   Pfam; PF00535; Glycos_transf_2; 1.
KW   Lipopolysaccharide biosynthesis; Transferase; Glycosyltransferase
SQ   SEQUENCE 257 AA; 29233 MW; D4DD7B57B002F990 CRC64;

Query Match   2.5%;   Score 7; DB 1;   Length 257;
Best Local Similarity 100.0%;   Pred. No. 62;
Matches   7;   Conservative   0; Mismatches   0; Indels   0; G

Qy   71 GLLLAVV 77
          |||||||
Db   229 GLLLAVV 235

RESULT 47
CN09_HUMAN
ID   CN09_HUMAN   STANDARD;   PRT;   277 AA.
AC   Q86T03; Q86U09; Q8WUC0; Q8WUG7; Q9NSU8;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Hypothetical protein C14orf9.
GN   C14ORF9.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC   TISSUE=Neuroblastoma, and T-cell;
RA   Li W.B., Gruber C., Jessee J., Polayes D.;
RT   "Full-length cDNA libraries and normalization.";
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 3).
RC   TISSUE=Testis;
RA   Ottenwaelder B., Obermaier B., Mewes H.-W., Gassenhuber J.,
RA   Wiemann S.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Brain, and Lung;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S
RA   Bosak S.A., McEwan K.J., McKernan K.J., Malek J.A., Gunaratne P.H
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- ALTERNATIVE PRODUCTS:
CC   Event=Alternative splicing; Named isoforms=3;
CC   Name=1;
CC   IsoId=Q86T03-1; Sequence=Displayed;
CC   Note=No experimental confirmation available;
CC   Name=2;
CC   IsoId=Q86T03-2; Sequence=VSP_007815;
CC   Note=No experimental confirmation available;

86T03-3; Sequence=VSP_007816, VSP_007817;
ty be due to intron retention.;

90; CAD61939.1; ALT_INIT.
25; CAD62347.1; ALT_INIT.
27; CAB70896.1; -.
57; AAH02867.2; -.
57; AAH20947.1; -.
T46382.
9299; C14orf9.
protein; Transmembrane; Alternative splicing.
POTENTIAL.
212   234   POTENTIAL.
244   266   POTENTIAL.
47   47   A -> AGKHAPPQ (in isoform 2).
                /FTId=VSP_007815.
48   224   APPFPEGHPAVLPGEDPPYSPLTSPDSGSAPMITCRVCQ
                SLINVEGKMHQHVVKCGVCNEATPIKNAPPGKKYVRCPCNC
                LLICKVTSQRIACRPYCKRIINLGPVHPGPLSPEPQPMGV
                RVICGHCKNTFLWTEFTDRTLARCPHCRKVSSIGRRYPRKR
                CICCFLLGLLLAV -> GKHAPPQGKFGRVRGAPRGTLKAG
                EGAGPRAEAGPSRQVRDCCTCDWARLPSLRNRDHSLGTEGG
                SEQPDRSANYEKPSELGQRVEDQKDFPTVEHQWGCK (in
                isoform 3).
                /FTId=VSP_007816.
225   277   Missing (in isoform 3).
                /FTId=VSP_007817.
208   208   P -> T (IN REF. 3; AAH20947).
77 AA; 29469 MW; A85FE1F736366CBC CRC64;

      2.5%;   Score 7;   DB 1;   Length 277;
arity 100.0%; Pred. No. 66;
onservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

LAV 76
|||
LAV 224

STANDARD;   PRT;   310 AA.
1;
(Rel. 41, Created)
(Rel. 43, Last sequence update)
(Rel. 43, Last annotation update)
beta (LT-beta) (Tumor necrosis factor C) (TNF-C) (Tumor
or ligand superfamily member 3).
OR TNFC.
(Woodchuck).
tazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
heria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;
95;
M.N.A.
4748; PubMed=10721723;
ll E.A., Brown C.L., Cullen J.M.;
ymphotoxin-alpha, -beta and tumor necrosis factor genes:
aracterization and biological activity.";
-305(2000).
: Cytokine that binds to LTBR/TNFRSF3. May play a specific
immune response regulation. Provides the membrane anchor
attachment of the heterotrimeric complex to the cell

CC   -!- SUBUNIT: Heterotrimer of either two LTB and one LTA subunits
CC   (less prevalent) two LTA and one LTB subunits (By similarity
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
CC   -!- SIMILARITY: Belongs to the tumor necrosis factor family.

DR   EMBL; AF096268; AAF34866.1; -.
DR   EMBL; AF095587; AAF34865.1; -.
DR   HSSP; P01374; 1TNR.
DR   InterPro; IPR006053; TNF_abc.
DR   InterPro; IPR006052; TNF_family.
DR   InterPro; IPR008983; TNF_like.
DR   InterPro; IPR003636; TNF_subf.
DR   Pfam; PF00229; TNF; 1.
DR   PRINTS; PR01234; TNECROSISFCT.
DR   ProDom; PD002012; TNF_subf; 1.
DR   SMART; SM00207; TNF; 1.
DR   PROSITE; PS00251; TNF_1; FALSE_NEG.
DR   PROSITE; PS50049; TNF_2; 1.
KW   Cytokine; Transmembrane; Glycoprotein; Signal-anchor.
FT   DOMAIN   1   27   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   28   48   SIGNAL-ANCHOR (TYPE-II MEMBRANE PRO
FT   (POTENTIAL).
FT   DOMAIN   49   310   EXTRACELLULAR (POTENTIAL).
FT   CARBOHYD 272   272   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CONFLICT 280   280   D -> H (IN REF. 1; AAF34865).
SQ   SEQUENCE   310 AA; 32644 MW; 73B354EFC8B3B3BE CRC64;

Query Match   2.5%; Score 7; DB 1; Length 310;
Best Local Similarity 100.0%; Pred. No. 73;
Matches   7; Conservative   0; Mismatches   0; Indels   0;

Qy   194 GLYLYLC 200
         |||||||
Db   181 GLYLYLC 187

RESULT 49
ISPH_XANCP
ID   ISPH_XANCP   STANDARD;   PRT;   316 AA.
AC   Q8PBG4;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   4-hydroxy-3-methylbut-2-enyl diphosphate reductase (EC 1.17.1.2)
GN   ISPH OR LYTB OR XCC1157.
OS   Xanthomonas campestris (pv. campestris).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xanthomonas.
OX   NCBI_TaxID=340;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 33913 / NCPPB 528;
RX   MEDLINE=22022145; PubMed=12024217;
RA   da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.
RA   Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N
RA   Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA   Camarote G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P
RA   Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H
RA   Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA   Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA   Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.
RA   Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.
RA   Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.
RA   Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V
RA   Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,

RA   Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA   Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA   de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA   Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA   Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA   de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA   da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA   da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA   de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuhako M.H.,
RA   Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA   Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT   "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL   Nature 406:151-159(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Temecula1 / ATCC 700964;
RX   MEDLINE=22421331; PubMed=12533478;
RA   Van Sluys M.A., de Oliveira M.C., Monteiro-Vitorello C.B.,
RA   Miyaki C.Y., Furlan L.R., Camargo L.E.A., da Silva A.C.R., Moon D.
RA   Takita M.A., Lemos E.G.M., Machado M.A., Ferro M.I.T., da Silva F.
RA   Goldman M.H.S., Goldman G.H., Lemos M.V.F., El-Dorry H., Tsai S.M.
RA   Carrer H., Carraro D.M., de Oliveira R.C., Nunes L.R., Siqueira W.
RA   Coutinho L.L., Kimura E.T., Ferro E.S., Harakava R., Kuramae E.E.,
RA   Marino C.L., Giglioti E., Abreu I.L., Alves L.M.C., do Amaral A.M.
RA   Baia G.S., Blanco S.R., Brito M.S., Cannavan F.S., Celestino A.V.,
RA   da Cunha A.F., Fenille R.C., Ferro J.A., Formighieri E.F., Kishi I
RA   Leoni S.G., Oliveira A.R., Rosa V.E. Jr., Sassaki F.T., Sena J.A.I
RA   de Souza A.A., Truffi D., Tsukumo F., Yanai G.M., Zaros L.G.,
RA   Civerolo E.L., Simpson A.J.G., Almeida N.F. Jr., Setubal J.C.,
RA   Kitajima J.P.;
RT   "Comparative analyses of the complete genome sequences of Pierce's
RT   disease and citrus variegated chlorosis strains of Xylella
RT   fastidiosa.";
RL   J. Bacteriol. 185:1018-1026(2003).
CC   -!- FUNCTION: Converts 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosph
CC   into isopentenyl diphosphate (IPP) and dimethylallyl diphospha
CC   (DMAPP) (By similarity).
CC   -!- CATALYTIC ACTIVITY: Isopentenyl diphosphate + NAD(P)(+) + H(2)
CC   (E)-4-hydroxy-3-methylbut-2-en-1-yl diphosphate + NAD(P)H.
CC   -!- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; seventh
CC   (last) step.
CC   -!- SIMILARITY: Belongs to the ispH family.
CC   -!-

DR   EMBL; AE004050; AAF85215.1; -.
DR   EMBL; AE012558; AAO29279.1; -.
DR   PIR; C82561; C82561.
DR   HAMAP; MF_00191; -; 1.
DR   InterPro; IPR003451; LytB.
DR   Pfam; PF02401; LYTB; 1.
DR   TIGRFAMs; TIGR00216; ispH_lytB; 1.
KW   Isoprene biosynthesis; Complete proteome; Oxidoreductase; NADP.
SQ   SEQUENCE   316 AA;   34704 MW; 1A2E80B9A98D334A CRC64;

Query Match   2.5%; Score 7; DB 1; Length 316;
Best Local Similarity   100.0%; Pred. No. 74;
Matches   7; Conservative   0; Mismatches   0; Indels   0; Ga

Qy   217 LVDGVLA 223
          |||||||
Db   276 LVDGVLA 282

RESULT 51
ODPB_BACSU
ID   ODPB_BACSU   STANDARD;   PRT;   324 AA.

, Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
antos M., Truffi D., Tsai S.M., White F.F.,
Kitajima J.P.;
the genomes of two Xanthomonas pathogens with differing
ties.";
-463(2002).
Converts 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate
ntenyl diphosphate (IPP) and dimethylallyl diphosphate
y similarity).
ACTIVITY: Isopentenyl diphosphate + NAD(P)(+) + H(2)O =
oxy-3-methylbut-2-en-1-yl diphosphate + NAD(P)H.
onmevalonate terpenoid biosynthesis pathway; seventh
p.
: Belongs to the ispH family.

T entry is copyright. It is produced through a collaboration
wiss Institute of Bioinformatics and the EMBL outstation -
ioinformatics Institute. There are no restrictions on its
rofit institutions as long as its content is in no way
his statement is not removed. Usage by and for commercial
res a license agreement (See http://www.isb-sib.ch/announce/
il to license@isb-sib.ch).

; AAM40456.1; -.
1; -; 1.
03451; LytB.
LYTB; 1.
00216; ispH_lytB; 1.
nthesis; Complete proteome; Oxidoreductase; NADP.
AA;   34641 MW; 6BF5A272A4CA1E4C CRC64;

rity   2.5%; Score 7; DB 1; Length 316;
nservative   0; Mismatches   0; Indels   0; Gaps   0;

LA 223
||
LA 282

STANDARD;   PRT;   316 AA.

el. 41, Created)
el. 41, Last sequence update)
el. 42, Last annotation update)
thylbut-2-enyl diphosphate reductase (EC 1.17.1.2).
R XF2416 OR PD1435.
iosa, and
iosa (strain Temecula1 / ATCC 700964).
eobacteria; Gammaproteobacteria; Xanthomonadales;
ae; Xylella.
1, 183190;

N.A.

17; PubMed=10910347;
, Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
onaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
, Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
anca S.C., Franco M.C., Frohme M., Furlan L.R.,
ldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
sel J.D., Junqueira M.L., Kemper B.L., Kitajima J.P.,
Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,

```
CC        lipoamide dehydrogenase (E3).
CC    -!- FUNCTION: THE B.SUBTILIS PDH COMPLEX POSSESSES ALSO BRANCHED-
CC        2-OXOACID DEHYDROGENASE (BCDH) ACTIVITY.
CC    -!- CATALYTIC ACTIVITY: Pyruvate + lipoamide = S-
CC        acetyldihydrolipoamide + CO(2).
CC    -!- COFACTOR: Thiamine pyrophosphate.
CC    -!- SUBUNIT: Heterodimer of an alpha and a beta chain.
CC    ---------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a col
CC    between the Swiss Institute of Bioinformatics and the EMBL out
CC    the European Bioinformatics Institute. There are no restrictio
CC    use by non-profit institutions as long as its content is ir
CC    modified and this statement is not removed. Usage by and for c
CC    entities requires a license agreement (See http://www.isb-sib.ch/
CC    or send an email to license@isb-sib.ch).
CC    ---------------------------------------------------------------
DR    EMBL; M57435; AAA62682.1; -.
DR    EMBL; AF012285; AAC24933.1; -.
DR    EMBL; Z99111; CAB13332.1; -.
DR    PIR; C36718; C36718.
DR    HSSP; P09061; 1QS0.
DR    SubtiList; BG10208; pdhB.
DR    InterPro; IPR009014; TransKeto_C_like.
DR    InterPro; IPR005476; Transketolase_C.
DR    InterPro; IPR005475; Transketolase_CR.
DR    Pfam; PF02779; transket_pyr; 1.
DR    Pfam; PF02780; transketolase_C; 1.
KW    Glycolysis; Oxidoreductase; Flavoprotein; Thiamine pyrophosphate;
KW    Complete proteome.
FT    INIT_MET    0    0       BY SIMILARITY.
SQ    SEQUENCE    324 AA;    35343 MW;  D2A7C9B32DED0F0D CRC64;

Query Match         2.5%;  Score 7;  DB 1;  Length 324;
Best Local Similarity  100.0%;  Pred. No. 75;
Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  C

Qy    59 LAIGLGL 65
          |||||||
Db    65 LAIGLGL 71


RESULT 52
SRA6_CAEEL
ID   SRA6_CAEEL    STANDARD;     PRT;    329 AA.
AC   Q09208;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Serpentine receptor class alpha 6 (Sra-6 protein).
GN   SRA-6 OR AH6.10.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoi
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Jassal B.;
RL   Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- SIMILARITY: BELONGS TO THE C.ELEGANS RECEPTOR-LIKE PROTEIN SI
CC       FAMILY.
CC    ---------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a col
CC    between the Swiss Institute of Bioinformatics and the EMBL out
CC    the European Bioinformatics Institute. There are no restrictio
CC    use by non-profit institutions as long as its content is il
CC    modified and this statement is not removed. Usage by and for c
CC    entities requires a license agreement (See http://www.isb-sib.ch).
CC    or send an email to license@isb-sib.ch).
CC    ---------------------------------------------------------------
DR    EMBL; Z48009; CAA88083.1; -.
DR    PIR; T18619; T18619.
```

```
DT   Rel. 18, Created)
DT   Rel. 18, Last sequence update)
DT   Rel. 42, Last annotation update)
DE   drogenase E1 component, beta subunit (EC 1.2.4.1) (S
DE   Da subunit).
GN   OR BSU14590.
OS   ilis.
OC   micutes; Bacillales; Bacillaceae; Bacillus.
OX   23;
RN   N.A.
RA
RL
RX   558; PubMed=1697575;
RA   lva A., Paulin L., Arvidson S., Palva I.;
RT   complex of Bacillus subtilis: sequence analysis and
RT   yruvate dehydrogenase.";
RL   172:5052-5063 (1990).
RN
RP   N.A.
RA
RL
RX   187; PubMed=8969500;
RA   aldwell R., Enfield L., Ferrari E.;
RT   E (124 degrees-127 degrees) region of the Bacillus
RT   chromosome: sequencing of a 27 kb segment and
RT   n of several genes in the area.";
RL   142:3033-3037(1996).
RN
RP   N.A.
RA
RA   , Ferrari E.;
RT   lysis of the mobA-amyS region of the Bacillus subtilis
RT
RL   L-1997) to the EMBL/GenBank/DDBJ databases.
RN
RP   N.A.
RA
RA   033; PubMed=9384377;
RA   sawara N., Moszer I., Albertini A.M., Alloni G.,
RA   tertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA   oursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA   Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA   dani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA   evine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA   Errington J., Fabret C., Ferrari E., Foulger D.,
RA   ita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA   aser P., Goffeau A., Golightly E.J., Grandi G.,
RA   Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA   holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA   amata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA   Koetter P., Koningstein G., Krogh S., Kumano M.,
RA   pidus A., Lardinois S., Lauber J., Lazarevic V.,
RA   ine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA   llado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA   eilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA   l T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Pujic C., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   volta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA   alan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA   acconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA   chiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA   nbutt R., Wedler E., Wedler H., Weitzenegger T.,
RA   ipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   oshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RA   a genome sequence of the Gram-positive bacterium Bacillus
RT   19-256(1997).
CC   The pyruvate dehydrogenase complex catalyzes the overall
CC   on of pyruvate to acetyl-CoA and CO(2). It contains
CC   copies of three enzymatic components: pyruvate
CC   nase (E1), dihydrolipoamide acetyltransferase (E2) and
```

```
RESULT 54
LEU3_THEAQ
ID   LEU3_THEAQ    STANDARD;    PRT;    344 AA.
AC   P24058;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   3-isopropylmalate dehydrogenase (EC 1.1.1.85) (Beta-IPM dehydroge
DE   (IMDH) (3-IPM-DH).
GN   LEUB.
OS   Thermus aquaticus.
OC   Bacteria; Deinococcus-Thermus; Deinococcus; Thermales; Thermaceae;
OC   Thermus.
OX   NCBI_TaxID=271;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YT1;
RX   MEDLINE=92041736; PubMed=1939005;
RA   Kirino H., Oshima T.;
RT   "Molecular cloning and nucleotide sequence of 3-isopropylmalate
RT   dehydrogenase gene (leuB) from an extreme thermophile, Thermus
RT   aquaticus YT-1.";
RL   J. Biochem. 109:852-857(1991).
CC   -!- FUNCTION: Catalyzes the oxidation of 3-carboxy-2-hydroxy-4-
CC       methylpentanoate (3-isopropylmalate) to 3-carboxy-4-methyl-2-
CC       oxopentanoate. The product decarboxylates to 4-methyl-2
CC       oxopentanoate.
CC   -!- CATALYTIC ACTIVITY: 3-carboxy-2-hydroxy-4-methylpentanoate +
CC       NAD(+) = 3-carboxy-4-methyl-2-oxopentanoate + NADH.
CC   -!- PATHWAY: Leucine biosynthesis; third step.
CC   -!- SUBUNIT: Homodimer.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: Belongs to the isocitrate and isopropylmalate
CC       dehydrogenases family. LeuB subfamily 1.
CC   -------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a coll
CC   between the Swiss Institute of Bioinformatics and the EMBL out
CC   the European Bioinformatics Institute. There are no restriction
CC   use by non-profit institutions as long as its content is in
CC   modified and this statement is not removed. Usage by and for c
CC   entries requires a license agreement (See http://www.isb-sib.ch/
CC   or send an email to license@isb-sib.ch).
CC   -------------------------------------------------------------------
DR   EMBL; D10700; BAA01542.1; -.
DR   PIR; S41223; DETWIT.
DR   HSSP; P00351; 1XAA.
DR   HAMAP; MF_01033; -; 1.
DR   InterPro; IPR001804; Isodh.
DR   InterPro; IPR004429; LeuB.
DR   Pfam; PF00180; isodh; 1.
DR   TIGRFAMs; TIGR00169; leuB; 1.
DR   PROSITE; PS00470; IDH_IMDH; 1.
KW   Oxidoreductase; Leucine biosynthesis; NAD.
SQ   SEQUENCE 344 AA; 36949 MW; 4F31A9444EB26408 CRC64;

Query Match     2.5%;  Score 7;  DB 1;  Length 344;
Best Local Similarity  100.0%;  Pred. No. 80;
Matches    7;  Conservative  0;  Mismatches  0;  Indels  0;  G

Qy    248 SGLLALR 254
              |||||||
Db     88 SGLLALR 94

RESULT 55
LEU3_DEIRA
ID   LEU3_DEIRA    STANDARD;    PRT;    352 AA.
AC   Q9RTH9;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   3-isopropylmalate dehydrogenase (EC 1.1.1.85) (Beta-IPM dehydroge
DE   (IMDH) (3-IPM-DH).
```

```
0; CE01451.
000344; Sra_chemrecept.
Sra; 1.
37; TMPROTEINSRA.
Multigene family.
26    46   POTENTIAL.
04   124   POTENTIAL.
13   163   POTENTIAL.
37   207   POTENTIAL.
38   258   POTENTIAL.
73   293   POTENTIAL.
9 AA;  37951 MW;  7C0963ADA53A29F3 CRC64;

rity    2.5%;  Score 7;  DB 1;  Length 329;
100.0%;  Pred. No. 76;
onservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

RR 140
   ||
RR 303

STANDARD;    PRT;    335 AA.

el. 38, Created)
el. 38, Last sequence update)
el. 38, Last annotation update)
associated protein.

cus (Rat).
azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
eria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
16;

N.A.
TISSUE=Uterus;
t A., Lessing J.B., Kraicer P.F., Kidron T.;
associated uterine protein.";
-1999) to the EMBL/GenBank/DDBJ databases.
R LOCATION: Integral membrane protein (Potential).
: Belongs to the OST3 family.

T entry is copyright. It is produced through a collaboration
Swiss Institute of Bioinformatics and the EMBL outstation -
Bioinformatics Institute. There are no restrictions on its
profit institutions as long as its content is in no way
his statement is not removed. Usage by and for commercial
res a license agreement (See http://www.isb-sib.ch/announce/
il to license@isb-sib.ch).

; AAB63294.2; -.
06B44; OST3_OST6.
06663; Thioredox_dom2.
OST3_OST6; 1.

6     26   POTENTIAL.
05   205   POTENTIAL.
0    230   POTENTIAL.
71   291   POTENTIAL.
21   321   POTENTIAL.
5 AA;  37992 MW;  AFB9DB35F1A06091 CRC64;

rity    2.5%;  Score 7;  DB 1;  Length 335;
100.0%;  Pred. No. 78;
onservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

AV 76
   ||
AV 198
```

5.
adiodurans.
nococcus-Thermus; Deinococci; Deinococcales;
e; Deinococcus.
99;

N.A.
TCC 13939 / DSM 20539 / NCIB 9279;
896; PubMed=10567266;
en J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
J., Lam P., McDonald L., Uterback T., Zalewski C.,
, Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,

nce of the radioresistant bacterium Deinococcus
1.";
571-1577(1999).
Catalyzes the oxidation of 3-carboxy-2-hydroxy-4-
tanoate (3-isopropylmalate) to 3-carboxy-4-methyl-2-
oate. The product decarboxylates to 4-methyl-2
oate.
ACTIVITY: 3-carboxy-2-hydroxy-4-methylpentanoate +
3-carboxy-4-methyl-2-oxopentanoate + NADH.
Leucine biosynthesis; third step.
Homodimer (By similarity).
AR LOCATION: Cytoplasmic (By similarity).
Y: Belongs to the isocitrate and isopropylmalate
nases family. LeuB subfamily 1.

OT entry is copyright. It is produced through a collaboration
Swiss Institute of Bioinformatics and the EMBL outstation -
Bioinformatics Institute. There are no restrictions on its
profit institutions as long as its content is in no way
this statement is not removed. Usage by and for commercial
ires a license agreement (See http://www.isb-sib.ch/announce/
ail to license@isb-sib.ch).

9; AAF11333.1; -.
G75355.
1XAA.

33; -; 1.
001804; Isodh.
004429; LeuB.
; isodh, 1.
R00169; leuB; 1.
470; IDH_IMDH; 1.
e; Leucine biosynthesis; NAD; Complete proteome.
2 AA; 37598 MW; 8BAE0E347F2AFA29 CRC64;

arity   2.5%; Pred. No. 81; Length 352;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

ALR 254
|||
ALR 94

STANDARD;   PRT;   357 AA.

Rel. 29, Created)
Rel. 29, Last sequence update)
Rel. 41, Last annotation update)
sphatase (EC 3.1.3.9) (G6Pase) (G-6-Pase).

(Mouse)
tazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
heria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94012716; PubMed=8407995;
RA Shelly L.L., Lei K.-J., Pan C.-J., Sakata S.F., Ruppert S.,
RA Schutz G., Chou J.Y.;
RT "Isolation of the gene for murine glucose-6-phosphatase, the enzy
RT deficient in glycogen storage disease type 1A.";
RL J. Biol. Chem. 268:21482-21485(1993).
RN [2]
RP SEQUENCE OF 1-76 FROM N.A.
RC STRAIN=129/Sv; TISSUE=Liver;
RX MEDLINE=97277298; PubMed=915220;
RA Streeper R.S., Svitek C.A., Chapman S., Greenbaum L.E., Taub R.,
RA O'Brien R.M.;
RT "A multicomponent insulin response sequence mediates a strong
RT repression of mouse glucose-6-phosphatase gene transcription by
RT insulin.";
RL J. Biol. Chem. 272:11698-11701(1997).
CC -!- FUNCTION: May be a single membrane channel protein acting bot
CC   a hydrolase and a translocase. It is the key enzyme in homeos
CC   regulation of blood glucose levels.
CC -!- CATALYTIC ACTIVITY: D-glucose 6-phosphate + H(2)O = D-glucose
CC   phosphate.
CC -!- PATHWAY: Gluconeogenesis and glycogenolysis; last step.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC   reticulum.
CC -!- TISSUE SPECIFICITY: Liver and kidney.
CC --
CC This SWISS-PROT entry is copyright. It is produced through a coll
CC between the Swiss Institute of Bioinformatics and the EMBL out
CC the European Bioinformatics Institute. There are no restrictio
CC use by non-profit institutions as long as its content is i
CC modified and this statement is not removed. Usage by and for c
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC --
DR EMBL; U00445; AAC52122.1; -.
DR EMBL; U91573; AAC53166.1; -.
DR PIR; A48589; A48589.
DR MGD; MGI:95607; G6pc.
DR InterPro; IPR008934; AcPase_VanPerase.
DR InterPro; IPR000326; PA_PTPase.
DR Pfam; PF01569; PAP2; 1.
DR SMART; SM00014; acidPPc; 1.
KW Glycogen biosynthesis; Hydrolase; Transmembrane; Glycoprotein;
KW Endoplasmic reticulum.
FT TRANSMEM   30   46   POTENTIAL.
FT TRANSMEM   59   75   POTENTIAL.
FT TRANSMEM   153   169   POTENTIAL.
FT TRANSMEM   211   227   POTENTIAL.
FT TRANSMEM   296   312   POTENTIAL.
FT TRANSMEM   333   349   POTENTIAL.
FT CARBOHYD   96   96   N-LINKED (GLCNAC...) (BY SIMILARIT
FT SITE   354   357   PREVENT SECRETION FROM ER (POTENTIAL
SQ SEQUENCE 357 AA; 40480 MW; 292F9FCB39582692 CRC64;

Query Match   2.5%;   Score 7; DB 1; Length 357;
Best Local Similarity 100.0%;   Pred. No. 82;
Matches   7;   Conservative   0;   Mismatches   0;   Indels   0;   G

Qy   61 LGLGLAL 67
         |||||||
Db   269 LGLGLAL 275

RESULT 57
G6PT_RAT
ID G6PT_RAT   STANDARD;   PRT;   357 AA.
AC P43428;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)

Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; G

Qy   61 LGLGLAI 67
               |||||||
Db   269 LGLGLAI 275

RESULT 58
PON1_RABIT
ID   PON1_RABIT   STANDARD;   PRT;   358 AA.
AC   P27170; Q9BGN1; Q9BGN2; Q9BGN3;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Serum paraoxonase/arylesterase 1 (EC 3.1.1.2)   (EC 3.1.8.1) (PON 1;
DE   (Serum aryldialkylphosphatase 1) (A-esterase 1) (Aromatic esterase
GN   PON1 OR PON.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Liver;
RX   MEDLINE=92031445; PubMed=1657140;
RA   Hassett C., Richter R.J., Humbert R., Chapline C., Crabb J.W.,
RA   Omiecinski C.J., Furlong C.E.;
RT   "Characterization of cDNA clones encoding rabbit and human serum
RT   paraoxonase: the mature protein retains its signal sequence.";
RL   Biochemistry 30:10141-10149(1991).
RN   [2]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX   MEDLINE=9345100; PubMed=8397745;
RA   Furlong C.E., Costa L.G., Hassett C., Richter R.J.,
RA   Sundstrom J.A., Adler D.A., Disteche C.M., Omiecinski C.J.,
RA   Chapline C., Crabb J.W.;
RT   "Human and rabbit paraoxonases: purification, cloning, sequencing
RT   mapping and role of polymorphism in organophosphate detoxificatio
RL   Chem. Biol. Interact. 87:35-48(1993).
RN   [3]
RP   SEQUENCE FROM N.A., FUNCTION, AND VARIANTS SER-81; GLU-92 AND GLY-
RC   STRAIN=New Zealand white; TISSUE=liver;
RX   MEDLINE=21163843; PubMed=11266077;
RA   Watson C.E., Draganov D.I., Billecke S.S., Bisgaier C.L., La Du B
RT   "Rabbits possess a serum paraoxonase polymorphism similar to the l
RT   Q192R.";
RL   Pharmacogenetics 11:123-134(2001).
RN   [4]
RP   CHARACTERIZATION, AND SEQUENCE OF 1-20.
RX   MEDLINE=92031444; PubMed=1718413;
RA   Furlong C.E., Richter R.J., Chapline C., Crabb J.W.;
RT   "Purification of rabbit and human serum paraoxonase.";
RL   Biochemistry 30:10133-10140(1991).
CC   -!- FUNCTION: Hydrolyzes the toxic metabolites of a variety of
CC   organophosphorus insecticides. Capable of hydrolyzing a broad
CC   spectrum of organophosphate substrates and a number of aromati
CC   carboxylic acid esters. Mediates an enzymatic protection of lc
CC   density lipoproteins against oxidative modification.
CC   -!- CATALYTIC ACTIVITY: Aryl dialkyl phosphate + H(2)O = dialkyl
CC   phosphate + an aryl alcohol.
CC   -!- CATALYTIC ACTIVITY: A phenyl acetate + H(2)O = a phenol + acet
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- PTM: Glycosylated.
CC   -!- PTM: The signal sequence is not cleaved.
CC   -!- POLYMORPHISM: There are two allelic forms, allozyme A and B, w
CC   differ in their substrate specificity. Both forms have similar
CC   arylesterase activity but allozyme B possesses greater paraoxc
CC   activity. Allozyme A is better at protecting LDL from oxidatic
CC   -!- SIMILARITY: Belongs to the paraoxonase family.
CC   -!- CAUTION: Ref.3 (AAK06398) sequence differs from that shown due
CC   a stop codon in position 355.

el. 41, Last annotation update)
phatase (EC 3.1.3.9) (G6Pase) (G-6-Pase).

cus (Rat).
azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
eria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
16;

N.A.

19; PubMed=7860767;
in S., Chuang E., Buikuisen W., Naji A., Taub R.A.;
f glucose-6-phosphatase gene and protein expression
ptive response in proliferating liver and diabetes.";
t. 95:832-841(1995).

N.A.

95; PubMed=8198588;
gaud D.M., El-Maghrabi M.R., Pan W., Subir M.,

a cDNA for the catalytic subunit of rat liver
phatase: regulation of gene expression in FAO hepatoma
in, dexamethasone and cAMP.";
ys. Res. Commun. 201:302-309(1994).

N.A.

-Dawley; TISSUE=Liver;
50; PubMed=8865366;
ajima H., Horikawa Y., Hamaguchi T., Yamasaki T.,
amba M., Hanafusa T., Matsuzawa Y.;
d distribution of glucose-6-phosphatase catalytic
ger RNA and its changes in the diabetic state.";
ol. Pathol. Pharmacol. 93:13-24(1996).
May be a single membrane channel protein acting both as
e and a translocase. It is the key enzyme in homeostatic
of blood glucose levels.
ACTIVITY: D-glucose 6-phosphate + H(2)O = D-glucose +

luconeogenesis and glycogenolysis; last step.
R LOCATION: Integral membrane protein. Endoplasmic

T entry is copyright. It is produced through a collaboration
wiss Institute of Bioinformatics and the EMBL outstation -
ioinformatics Institute. There are no restrictions on its
rofit institutions as long as its content is in no way
his statement is not removed. Usage by and for commercial
res a license agreement. (See http://www.isb-sib.ch/announce/
il to license@isb-sib.ch).

AAA74381.1; ALT_INIT.
AAA19966.1; -.
BAA24348.1; -.
C2371.
08934; AcPase VanPerase.
00326; PA_PTPase.
PAP2; 1.
; acidPPc; 1.
nthesis; Hydrolase; Transmembrane; Glycoprotein;
ticulum.
0   46   POTENTIAL.
9   75   POTENTIAL.
3   169   POTENTIAL.
1   227   POTENTIAL.
6   312   POTENTIAL.
3   349   POTENTIAL.
6   96   N-LINKED (GLCNAC...)   (BY SIMILARITY).
4   357   PREVENT SECRETION FROM ER (POTENTIAL).
8   118   G -> V (IN REF. 2).
' AA, 40555 MW; C4496DE102F4244D CRC64;

    2.5%; Score 7; DB 1; Length 357;

```
OT entry is copyright. It is produced through a collaboration
Swiss Institute of Bioinformatics and the EMBL outstation -
Bioinformatics Institute. There are no restrictions on its
profit institutions as long as its content is in no way
this statement is not removed. Usage by and for commercial
ires a license agreement (See http://www.isb-sib.ch/announce/
ail to license@isb-sib.ch).

AAA31452.1; -.
AAB27713.2; -.
1; AAK06398.1; ALT_TERM.
2; AAK06399.1; -.
3; AAK06400.1; -.
B40354.
6; C:extracellular; NAS.
9; F:antioxidant activity; IDA.
3; F:aryldialkylphosphatase activity; NAS.
4; F:arylesterase activity; IDA.
3; P:response to organophosphorus; IDA.
002640; Arylesterase.
008363; Paraoxonase1.
008364; Paraoxonase2.
; Arylesterase; 1.
85; PARAOXONASE.
86; PARAOXONASE1.
87; PARAOXONASE2.
tioxidant; Glycoprotein; Signal; Multigene family;

0       0      NOT CLEAVED.
1       ?      BY SIMILARITY.
41    352
49     49      N-LINKED (GLCNAC...  )   (POTENTIAL).
52    252      N-LINKED (GLCNAC...  )   (POTENTIAL).
69    269      N-LINKED (GLCNAC...  )   (POTENTIAL).
23    323      N-LINKED (GLCNAC...  )   (POTENTIAL).
81     81      P -> S  (IN ALLELE A).
92     92      K -> E  (IN ALLELE A).
00    100      S -> G  (IN ALLELE A).
66     66      A -> S  (IN REF. 3).
19    319      A -> V  (IN REF. 3; AAK06398).
8 AA;  39878 MW;  C40C44E5F6E5EFDF CRC64;

arity        2.5%;  Score 7;  DB 1;  Length 358;
100.0%;  Pred. No. 82;  Mismatches 0;  Indels 0;  Gaps 0;
onservative 0;  Mismatches 0;  Indels

LAL 67
|||
LAL 15
```

```
STANDARD;     PRT;    365 AA.

Rel. 35, Created)
Rel. 35, Last sequence update)
Rel. 43, Last annotation update)
oxidoreductase chain 8 (EC 1.6.99.5) (NADH dehydrogenase
NDH-1, chain 8).
ophilus.
nococcus-Thermus; Deinococci; Thermales; Thermaceae;
4;
N.A.
ATCC 27634;
490; PubMed=9020134;
S.S., Sled' V.D., Ohnishi T., Yagi T.;
ranslocating NADH-quinone oxidoreductase (NDH-1) of
bacterium Thermus thermophilus HB-8. Complete DNA
```

```
RT   sequence of the gene cluster and thermostable properties of the
RT   expressed NQO2 subunit.";
RL   J. Biol. Chem. 272:4201-4211(1997).
CC   -!- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iro
CC       sulfur (Fe-S) centers, to quinones in the respiratory chain.
CC       immediate electron acceptor for the enzyme in this species is
CC       believed to be menaquinone. Couples the redox reaction to pro
CC       translocation (for every two electrons transferred, four hydr
CC       ions are translocated across the cytoplasmic membrane), and t
CC       conserves the redox energy in a proton gradient.
CC   -!- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.
CC   -!- SUBUNIT: COMPOSED OF 14 DIFFERENT SUBUNITS. SUBUNITS NQO7-14
CC       CONSTITUTE THE MEMBRANE SECTOR OF THE COMPLEX.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to the complex I subunit 1 family.
CC
CC   This SWISS-PROT entry is copyright. It is produced through a coll
CC   between the Swiss Institute of Bioinformatics and the EMBL out
CC   the European Bioinformatics Institute. There are no restriction
CC   use by non-profit institutions as long as its content is in
CC   modified and this statement is not removed. Usage by and for c
CC   entities requires a license agreement (See http://www.isb-sib.ch/
CC   or send an email to license@isb-sib.ch).
CC
DR   EMBL; U52917; AAA97945.1; -.
DR   PIR; T11905; T11905.
DR   InterPro; IPR001694; Resp_NADH_dh1.
DR   Pfam; PF00146; NADHdh; 1.
DR   PROSITE; PS00667; COMPLEX1_ND1_1; 1.
DR   PROSITE; PS00668; COMPLEX1_ND1_2; 1.
KW   Oxidoreductase; NAD; Quinone; Transmembrane.
FT   TRANSMEM     11     31       POTENTIAL.
FT   TRANSMEM     80    100       POTENTIAL.
FT   TRANSMEM    120    140       POTENTIAL.
FT   TRANSMEM    157    177       POTENTIAL.
FT   TRANSMEM    192    212       POTENTIAL.
FT   TRANSMEM    252    272       POTENTIAL.
FT   TRANSMEM    273    293       POTENTIAL.
FT   TRANSMEM    310    330       POTENTIAL.
FT   TRANSMEM    336    356       POTENTIAL.
SQ   SEQUENCE   365 AA;  41008 MW;  AE920CC029333C09 CRC64;

Query Match      2.5%;  Score 7;  DB 1;  Length 365;
Best Local Similarity  100.0%;  Pred. No. 84;
Matches   7; Conservative 0;  Mismatches 0;  Indels 0;  C

Qy      61 LGLGLAL 67
           |||||||
Db     164 LGLGLAL 170

RESULT 60
BENE_ACICA
ID   BENE_ACICA    STANDARD;    PRT;   394 AA.
AC   P07775;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Benzoate membrane transport protein.
GN   BENE.
OS   Acinetobacter calcoaceticus.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Moraxellaceae; Acinetobacter.
OX   NCBI_TaxID=471;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BD413 / ADP1;
RX   MEDLINE=91358314; PubMed=1885518;
RA   Neidle E.L., Hartnett C., Ornston N.L., Bairoch A., Rekik M.,
RA   Harayama S.;
RT   "Nucleotide sequences of the Acinetobacter calcoaceticus benABC
RT   for benzoate 1,2-dioxygenase reveal evolutionary relationships an
RT   multicomponent oxygenases.";
```

173:5385-5395(1991).
PROBABLY INVOLVED IN THE TRANSPORT OF BENZOATE.
R LOCATION: Integral membrane protein.
OUS: THE BEN OPERON ENCODE THE PROTEINS RESPONSIBLE FOR
ATION OF BENZOATE TO CATECHOL.

; AAC46440.1; -
23481.
04711; BenE.
  BenE; 1.
53; BenE; 1.
00843; benE; 1.
matic hydrocarbons catabolism; Transmembrane.
  AA; 42288 MW; 5466C4D460784BB5 CRC64;

rity 2.5%; Score 7; DB 1; Length 394;
 100.0%; Pred. No. 90;
nservative 0; Mismatches 0; Indels 0; Gaps 0;

LV 57
||
LV 88

STANDARD;   PRT;   396 AA.

el. 38, Created)
el. 38, Last sequence update)
el. 42, Last annotation update)
g protein precursor 1 (DHH-1) (Cephalic hedgehog
H).

(African clawed frog).
azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
achia; Anura; Mesobatrachia; Pipoidea; Pipidae;
enopus.
5;

N.A.

52; PubMed=7671800;
Grew L.L., Lai C.-J., Lee J.J., von Kessler D.P.,
chy P.A.;
ession and shared activities of members of the hedgehog
  Xenopus laevis.";
1:2337-2347(1995).
SIGNAL INVOLVED IN THE EARLY INDUCTION AND PATTERNING OF
AL ECTODERM, NERVOUS SYSTEM AND SOMITES. INDUCES ECTOPIC
ND FORMATION IN EMBRYOS.
R LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
E THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
CE. HEDGEHOG PROTEIN IS ALSO SECRETED IN EITHER CLEAVED
ED FORM TO MEDIATE SIGNALING TO OTHER CELLS (BY
).
ITAL STAGE: DETECTABLE WITHIN THE EARLY GASTRULA. IN THE
PRESSION BECOMES RESTRICTED TO ANTERIOR STRUCTURES,
NG BOTH NEURAL PLATE AND ENDODERMAL CELLS.
-terminal domain displays an autoproteolysis activity
esterol transferase activity. Both activities result in
ge of the full-length protein and covalent attachment of
rol moiety to the C-terminal of the newly generated N-
ragment (N-product). This covalent modification appears
essential role in restricting the spatial distribution

CC   of the protein activity to the cell surface. The N-product is
CC   active species in both local and long-range signaling, wherea:
CC   C-product has no signaling activity (By similarity).
CC -!- SIMILARITY: Belongs to the hedgehog family.
CC
DR   EMBL; U26349; AAA85163.1; -.
DR   HSSP; Q62226; 1VHH.
DR   MEROPS; C46.UPW; -.
DR   InterPro; IPR009045; Hedgehog/DD_pept.
DR   InterPro; IPR003587; Hedgehog_hint_N.
DR   InterPro; IPR003586; Hedgehog_hintC.
DR   InterPro; IPR000320; HH_signal.
DR   InterPro; IPR006141; Intein_S.
DR   InterPro; IPR001767; Pept_C46_hint.
DR   InterPro; IPR001657; Peptidase_C46.
DR   Pfam; PF01085; HH_signal; 1.
DR   Pfam; PF01079; HinT; 1.
DR   PRINTS; PR00632; SONICHHOG.
DR   ProDom; PD003042; HH_signal; 1.
DR   SMART; SM00305; HintC; 1.
DR   SMART; SM00306; HintN; 1.
DR   PROSITE; PS50817; INTEIN_N_TER; 1.
KW   Developmental protein; Autocatalytic cleavage; Hydrolase; Proteas
KW   Signal; Lipoprotein; Palmitate.
FT   SIGNAL   1   22   POTENTIAL.
FT   CHAIN   23   396   DESERT HEDGEHOG PROTEIN 1.
FT   CHAIN   23   197   DESERT HEDGEHOG PROTEIN 1 N-PRODUCT.
FT   CHAIN   198   396   DESERT HEDGEHOG PROTEIN 1 C-PRODUCT.
FT   DOMAIN   276   279   POLY-SER.
FT   SITE   197   198   CLEAVAGE (AUTO-) (BY SIMILARITY).
FT   SITE   267   267   INVOLVED IN AUTO-CLEAVAGE (BY
FT   SIMILARITY).
FT   ACT_SITE 270   270   ESSENTIAL FOR AUTO-CLEAVAGE (BY
FT   SIMILARITY).
FT   LIPID   23   23   N-palmitoyl cysteine (By similarity)
FT   LIPID   197   197   Cholesterol glycine ester (By
FT   similarity).
SQ   SEQUENCE 396 AA; 44087 MW; 774A3EC2268A5EE9 CRC64;

Query Match   2.5%; Score 7; DB 1; Length 396;
Best Local Similarity 100.0%; Pred. No. 90;
Matches   7; Conservative 0; Mismatches 0; Indels 0; G

Qy   216 LLVDGVL 222
        |||||||
Db   331 LLVDGVL 337

RESULT 62
DHH2_XENLA
ID   DHH2_XENLA   STANDARD;   PRT;   398 AA.
AC   Q91611;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Desert hedgehog protein precursor 2 (DHH-2) (Hedgehog protein 4) (
DE   HH4).
GN   HH4.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.

; PubMed=7671800;
cGrew L.L., Lai C.-J., Lee J.J., von Kessler D.P.,
achy P.A.;
ression and shared activities of members of the hedgehog
of Xenopus laevis.";
21:2337-2347(1995).
SIGNAL INVOLVED IN THE EARLY INDUCTION AND PATTERNING OF
SAL ECTODERM, NERVOUS SYSTEM AND SOMITES. INDUCES ECTOPIC
AND FORMATION IN EMBRYOS.
AR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
LE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
ACE. HEDGEHOG PROTEIN IS ALSO SECRETED IN EITHER CLEAVED
VED FORM TO MEDIATE SIGNALING TO OTHER CELLS (BY
Y).
C-terminal domain displays an autoproteolysis activity
lesterol transferase activity. Both activities result in
rage of the full-length protein and covalent attachment of
erol moiety to the C-terminal of the newly generated N-
fragment (N-product). This covalent modification appears
n essential role in restricting the spatial distribution
otein activity to the cell surface. The N-product is the
ecies in both local and long-range signaling, whereas the
has no signaling activity (By similarity).
Y: Belongs to the hedgehog family.

. AAA85164.1; -.
.1VHH.
IPW; -.
009045; Hedgehog/DD_pept.
003587; Hedgehog_hint_N.
003586; Hedgehog_hintC.
000320; HH_signal.
001767; Pept_C46_hint.
001657; Peptidase_C46.
; HH_signal; 1.
); Hint; 1.
532; SONICHHOG.
042; HH_signal; 1.
05; HintC; 1.
06; HintN; 1.

protein; Autocatalytic cleavage; Hydrolase; Protease;
protein; Palmitate.
1   23   POTENTIAL.
24   398   DESERT HEDGEHOG PROTEIN 2.
24   199   DESERT HEDGEHOG PROTEIN 2 N-PRODUCT.
200   398   DESERT HEDGEHOG PROTEIN 2 C-PRODUCT.
278   281   POLY-SER.
199   200   CLEAVAGE (AUTO-) (BY SIMILARITY).
269   269   INVOLVED IN AUTO-CLEAVAGE (BY
              SIMILARITY).
272   272   ESSENTIAL FOR AUTO-CLEAVAGE (BY
              SIMILARITY).
24   24   N-palmitoyl cysteine (By similarity).
199   199   Cholesterol glycine ester (By
              similarity).

98 AA; 44458 MW; DBC23AF85F69DD08 CRC64;

larity   2.5%;   Score 7;   DB 1;   Length 398;
Conservative 100.0%; Pred. No. 90;
Conservative   0; Mismatches   0; Indels   0; Gaps   0;

DGVL 222
|||
DGVL 339

RESULT 63
SELP_BOVIN
ID   SELP_BOVIN   STANDARD;   PRT;   402 AA.
AC   P49907; O19003;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Selenoprotein P-like protein precursor.
GN   SEPP1 OR SELP.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Cerebellum;
RX   MEDLINE=95364621; PubMed=7637580;
RA   Saijoh K., Saito N., Lee M.J., Fujii M., Kobayashi T., Sumino K.;
RT   "Molecular cloning of cDNA encoding a bovine selenoprotein P-like
RT   protein containing 12 selenocysteines and a (His-Pro) rich domain
RT   insertion, and its regional expression.";
RL   Brain Res. Mol. Brain Res. 30:301-311(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98019090; PubMed=9358058;
RA   Fujii M., Saijoh K., Kobayashi T., Fujii S., Lee M.J., Sumino K.;
RT   "Analysis of bovine selenoprotein P-like protein gene and availa
RT   of metal responsive element (MRE) located in its promoter.";
RL   Gene 193:211-217(1997).
CC   -!- FUNCTION: It constitutes a major selenium pool in the brain a
CC   may play an important role in developing and/or modulating t
CC   morphology of neurons and/or glial cells.
CC   -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -!- TISSUE SPECIFICITY: Brain and kidney. Most prominently expre
CC   in the cerebellar cortex, hippocampus and olfactory bulb.
CC   -!- MISCELLANEOUS: The selenocysteines are all encoded by the op
CC   codon, UGA.
CC   ---
CC   ---
DR   EMBL; D25220; BAA04949.2; -.
DR   EMBL; D88033; BAA23414.1; -.
DR   EMBL; D88031; BAA23414.1; JOINED.
DR   EMBL; D88032; BAA23414.1; JOINED.
DR   InterPro; IPR007672; SelP_C.
DR   InterPro; IPR007671; SelP_N.
DR   Pfam; PF04593; SelP_C; 1.
DR   Pfam; PF04592; SelP_N; 1.
KW   Glycoprotein; Signal; Selenium; Selenocysteine; Repeat.
FT   SIGNAL   1   19
FT   CHAIN   20   402   BY SIMILARITY
FT   SELENOPROTEIN P-LIKE PROTEIN.
FT   SE_CYS   59   59
FT   SE_CYS   297   297
FT   SE_CYS   307   307
FT   SE_CYS   338   338
FT   SE_CYS   350   350
FT   SE_CYS   363   363
FT   SE_CYS   365   365
FT   SE_CYS   372   372
FT   SE_CYS   388   388
FT   SE_CYS   390   390
FT   SE_CYS   397   397
FT   SE_CYS   399   399
FT   DOMAIN   204   239   H-P REPEATS.

```
         266    POLY-HIS
         181    SRPQ -> KALE (IN REF. 2).
         256    T -> P (IN REF. 2).
         282    L -> V (IN REF. 2).
         312    Y -> D (IN REF. 2).
      AA; 45018 MW; B7CF18751FB0E8FF CRC64;

        2.5%; Score 7; DB 1; Length 402;
   100.0%; Pred. No. 91;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

LA 68
LA 10
```

```
PGKC_ALCEU
ID   PGKC_ALCEU          STANDARD;      PRT;    413 AA.
AC   P50319;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Phosphoglycerate kinase, chromosomal (EC 2.7.2.3).
GN   CBBKC.
OS   Alcaligenes eutrophus (Ralstonia eutropha).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Burkholderiaceae; Ralstonia.
OX   NCBI_TaxID=510;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H16 / DSM 428 / ATCC 17699;
RX   MEDLINE=9528415; PubMed=7763137;
RA   Schaeferfohann J., Yoo J.-G., Bowien B.;
RT   "Analysis of the genes forming the distal parts of the two cbb CO:
RT   fixation operons from Alcaligenes eutrophus.";
RL   Arch. Microbiol. 163:291-299(1995).
CC   -!- CATALYTIC ACTIVITY: ATP + 3-phospho-D-glycerate = ADP + 3-
CC       phospho-D-glyceroyl phosphate.
CC   -!- PATHWAY: Calvin cycle.
CC   -!- SUBUNIT: Monomer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: Belongs to the phosphoglycerate kinase family.
CC   ------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a coll
CC   between the Swiss Institute of Bioinformatics and the EMBL out
CC   the European Bioinformatics Institute. There are no restriction
CC   use by non-profit institutions as long as its content is in
CC   modified and this statement is not removed. Usage by and for c
CC   entities requires a license agreement (See http://www.isb-sib.ch/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------
DR   EMBL; U12422; AAC43444.1; -.
DR   PIR; I39551; I39551.
DR   HSSP; P18912; 1PHP.
DR   HAMAP; MF_00145; -; 1.
DR   InterPro; IPR001576; PGK.
DR   Pfam; PF00162; PGK; 1.
DR   PRINTS; PR00477; PHGLYCKINASE.
DR   PROSITE; PS00111; PGLYCERATE_KINASE; 1.
KW   Transferase; Kinase; Calvin cycle.
SQ   SEQUENCE   413 AA;  42283 MW;  6B4C9D195566A90D CRC64;

Query Match          2.5%;  Score 7;  DB 1;  Length 413;
Best Local Similarity  100.0%;  Pred. No. 93;
Matches    7;  Conservative   0;  Mismatches   0;  Indels   0;  G

Qy    229 EFSATAA 235
              |||||||
Db    288 EFSATAA 294
```

```
RESULT 66
NH59_CAEEL
ID   NH59_CAEEL          STANDARD;      PRT;    416 AA.
AC   Q9TXU1; Q9GTF2;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Nuclear hormone receptor family member nhr-59.
GN   NHR-59 OR T27B7.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoid
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bogan A., Maina C.V., Yamamoto K., Cohen P., Sluder A.E.;
RT   "Caenorhabditis elegans nuclear receptor sequences exhibit biophy
RT   compatibility with the ligand-binding domain fold.";
```

```
                STANDARD;      PRT;    412 AA.

el. 34, Created)
el. 34, Last sequence update)
el. 41, Last annotation update)
te kinase, plasmid (EC 2.7.2.3).

trophus (Ralstonia eutropha).
asmid pHG1.
eobacteria; Betaproteobacteria; Burkholderiales;
ae; Ralstonia.

N.A.
DSM 428 / ATCC 17699;
15; PubMed=7763137;
J., Yoo J.-G., Bowien B.;
he genes forming the distal parts of the two cbb CO2
ns from Alcaligenes eutrophus.";
l. 163:291-299(1995).
ACTIVITY: ATP + 3-phospho-D-glycerate = ADP + 3-
glyceroyl phosphate.
alvin cycle.
onomer (By similarity).
R LOCATION: Cytoplasmic.
: Belongs to the phosphoglycerate kinase family.
-----------------------------------------------
T entry is copyright. It is produced through a collaboration
wiss Institute of Bioinformatics and the EMBL outstation -
bioinformatics Institute. There are no restrictions on its
rofit institutions as long as its content is in no way
his statement is not removed. Usage by and for commercial
res a license agreement (See http://www.isb-sib.ch/announce/
il to license@isb-sib.ch).
-----------------------------------------------
AAC43447.1; -.
1PHP.
45; -; 1.
001576; PGK.
77; PGK; 1.
11; PGLYCKINASE.
11; PGLYCERATE_KINASE; 1.
Kinase; Calvin Cycle; Plasmid.
2 AA;  42298 MW;  9SB4C666859E7274 CRC64;

arity        2.5%;  Score 7;  DB 1;  Length 412;
   100.0%;  Pred. No. 93;
onservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

AAA 235
CAA 293
```

RX   MEDLINE=90306038; PubMed=1694763;
RA   Pages G., Rouayrenc J.F., le Cam G., Mariller M., le Cam A.;
RT   "Molecular characterization of three rat liver serine-protease
RT   inhibitors affected by inflammation and hypophysectomy. Protein a
RL   mRNA analysis and cDNA cloning.";
RL   Eur. J. Biochem. 190:385-391(1990).
RN   [3]
RP   SEQUENCE OF 203-408 FROM N.A.
RX   MEDLINE=87144617; PubMed=3493437;
RA   Hill R.E., Hastie N.D.;
RT   "Accelerated evolution in the reactive centre regions of serine
RT   protease inhibitors.";
RL   Nature 326:96-99(1987).
CC   -!- FUNCTION: INHIBITS TRYPSIN, BUT NOT CHYMOTRYPSIN OR
CC   ELASTASE.
CC   -!- TISSUE SPECIFICITY: Liver.
CC   -!- INDUCTION: By acute inflammation.
CC   -!- SIMILARITY: Belongs to the serpin family.
CC   -!- CAUTION: It is uncertain whether Met-1 or Met-11 is the initi

DR   EMBL; D00753; BAA00650.1; -.
DR   EMBL; X16359; CAA34408.1; -.
DR   EMBL; X13150; CAA31548.1; -.
DR   HSSP; P01011; 2ACH.
DR   InterPro; IPR000215; Serpin.
DR   Pfam; PF00079; serpin; 1.
DR   SMART; SM00093; SERPIN; 1.
DR   PROSITE; PS00284; SERPIN; 1.
KW   Serpin; Serine protease inhibitor; Glycoprotein; Signal.
FT   SIGNAL   1   29   BY SIMILARITY.
FT   CHAIN   30   418   CONTRAPSIN-LIKE PROTEASE INHIBITOR (
FT   ACT_SITE   381   382   REACTIVE BOND (POTENTIAL).
FT   CARBOHYD   104   104   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   258   258   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   269   269   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CONFLICT   12   12   A -> D (IN REF. 2).
FT   CONFLICT   20   20   C -> S (IN REF. 2).
FT   CONFLICT   68   68   D -> H (IN REF. 2).
FT   CONFLICT   92   92   N -> S (IN REF. 2).
FT   CONFLICT   236   236   P -> S (IN REF. 2).
FT   CONFLICT   323   323   E -> EE (IN REF. 3).
FT   CONFLICT   334   334   V -> A (IN REF. 3).
FT   CONFLICT   353   353   S -> F (IN REF. 3).
FT   CONFLICT   354   354   Q -> E (IN REF. 3).
FT   CONFLICT   381   381   M -> V (IN REF. 3).
SQ   SEQUENCE   418 AA; 46651 MW; AADEFF087190B44F CRC64;

Query Match   2.5%; Score 7; DB 1; Length 418;
Best Local Similarity   100.0%; Pred. No. 94;
Matches   7; Conservative   0; Mismatches   0; Indels   0; C

Qy   74 LAVVSLG 80
           |||||||
Db   82 LAVVSLG 88

RESULT 68
ENO_PYRAE
ID   ENO_PYRAE   STANDARD;   PRT;   419 AA.
AC   Q8ZYE7;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho
DE   glycerate hydro-lyase).

Y-2000) to the EMBL/GenBank/DDBJ databases.
N.A.
1 N2;
sley P., O'Brien D.;
G-1997) to the EMBL/GenBank/DDBJ databases.
N-2002) to the EMBL/GenBank/DDBJ databases.
Orphan nuclear receptor.
AR LOCATION: Nuclear (Potential).
Y: Belongs to the nuclear hormone receptor family.

7; AAG15156.1; -.
5; AAF02172.2; -.
1A6Y.
7.1; CE24017.
008946; Str_ncl_receptor.
001628; Znf_C4steroid.
; zf-C4; 1.
47; STROIDFINGER.
035; Znf C4steroid; 1.
9; ZnF_C4; 1.
031; NUCLEAR RECEPTOR; FALSE NEG.
nscription regulation; DNA-binding; Nuclear protein;
20   87   NUCLEAR RECEPTOR-TYPE.
20   40   C4-TYPE.
57   82   C4-TYPE.
6 AA; 47721 MW; D4A21CA587ED96D3 CRC64;

arity 2.5%; Score 7; DB 1; Length 416;
       100.0%; Pred. No. 94;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

ISA 10
  ||
ISA 177

      STANDARD;   PRT;   418 AA.
2;
Rel. 09, Created)
Rel. 35, Last sequence update)
Rel. 35, Last annotation update)
ke protease inhibitor 6 precursor (CPI-26) (Serine
bitor 3) (SPI-3) (SPI-2.2).
icus (Rat).
tazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
heria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
)116;
1 N.A.
1305; PubMed=1864837;
ata S., Misumi Y., Takami N., Ikehara Y.;
oning and characterization of rat contrapsin-like
bitor and related proteins.";
09:243-250(1991).
.1-408 FROM N.A.

```
RA   Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA   Chandler M., Choisne N., Claudel-Renard C., Cunnac S., Demange N.,
RA   Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA   Siguier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA   Weissenbach J., Boucher C.A.;
RT   "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL   Nature 415:497-502(2002).
CC   -!- CATALYTIC ACTIVITY: ATP + 3-phospho-D-glycerate = ADP + 3-
CC       phospho-D-glyceroyl phosphate.
CC   -!- PATHWAY: Second phase of glycolysis; second step.
CC   -!- SUBUNIT: Monomer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC   -!- SIMILARITY: Belongs to the phosphoglycerate kinase family.
CC   --------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a colla
CC   between the Swiss Institute of Bioinformatics and the EMBL outs
CC   the European Bioinformatics Institute. There are no restriction
CC   use by non-profit institutions as long as its content is in
CC   modified and this statement is not removed. Usage by and for co
CC   entities requires a license agreement (See http://www.isb-sib.ch/a
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------
DR   EMBL; AL646060; CAD14101.1; -.
DR   HAMAP; MF_00145; -; 1.
DR   InterPro; IPR001576; PGK.
DR   Pfam; PF00162; PGK; 1.
DR   PRINTS; PR00477; PHGLYCKINASE.
DR   PROSITE; PS00111; PGLYCERATE_KINASE; 1.
KW   Transferase; Kinase; Glycolysis; Complete proteome.
SQ   SEQUENCE   419 AA;  43268 MW;  8192410852B3D81 CRC64;

Query Match   2.5%;  Score 7;  DB 1;  Length 419;
Best Local Similarity   100.0%;  Pred.No. 95;
Matches   7;  Conservative   0;  Mismatches   0;  Indels   0;  G

Qy   229 EFSATAA 235
          |||||||
Db   289 EFSATAA 295

RESULT 70
YJ54_YEAST
ID   YJ54_YEAST   STANDARD;   PRT;   423 AA.
AC   P47130;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   Hypothetical 49.5 kDa protein in MIR1-STE18 intergenic region.
GN   YJR084W OR J1860.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C;
RX   MEDLINE=96437976; PubMed=8840504;
RA   Huang M.E., Manus V., Chuat J.-C., Galibert F.;
RT   "Analysis of a 62 kb DNA sequence of chromosome X reveals 36 open
RT   reading frames and a gene cluster with a counterpart on chromosom
RT   XI.";
RL   Yeast 12:869-875(1996).
CC   --------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a colla
CC   between the Swiss Institute of Bioinformatics and the EMBL outs
CC   the European Bioinformatics Institute. There are no restriction
CC   use by non-profit institutions as long as its content is in
CC   modified and this statement is not removed. Usage by and for co
CC   entities requires a license agreement (See http://www.isb-sib.ch).
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------
DR   EMBL; Z49584; CAA89611.1; -.
DR   EMBL; L47993; AAB33307.1; -.
```

```
.rophilum.
rchaeota; Thermoprotei; Thermoproteales;
ae; Pyrobaculum.
73;
N.A.
TCC 51768 / DSM 7523;
97; PubMed=11792869;
T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
ce of the hyperthermophilic crenarchaeon Pyrobaculum
ad. Sci. U.S.A. 99:984-989(2002).
ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
Magnesium is required for catalysis and for stabilizing
(By similarity).
lycolysis.
onomer (By similarity).
R LOCATION: Cytoplasmic (By similarity).
: Belongs to the enolase family.
--------------------------------------------------------------
T entry is copyright. It is produced through a collaboration
wiss Institute of Bioinformatics and the EMBL outstation -
Bioinformatics Institute. There are no restrictions on its
profit institutions as long as its content is in no way
his statement is not removed. Usage by and for commercial
res a license agreement (See http://www.isb-sib.ch/announce/
il to license@isb-sib.ch).
--------------------------------------------------------------
'; AAL63046.1; -.
8; -.; 1
00941; Enolase.
enolase; 1.
enolase_N; 1.
8; ENOLASE.
02; Enolase; 1.
64; ENOLASE; FALSE_NEG.
sis; Magnesium; Complete proteome.
.1   151   BY SIMILARITY.
.0   240   MAGNESIUM (BY SIMILARITY).
.3   283   MAGNESIUM (BY SIMILARITY).
.9   309   MAGNESIUM (BY SIMILARITY).
, AA;  45377 MW;  1B6068A3137D7F9E CRC64;

rity   2.5%;  Score 7;  DB 1;  Length 419;
100.0%;  Pred.No. 95;
nservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

LP 16
||
LP 141

STANDARD;   PRT;   419 AA.
el. 41, Created)
el. 41, Last sequence update)
el. 41, Last annotation update)
te kinase (EC 2.7.2.3).
OR RS04894.
nacearum (Pseudomonas solanacearum).
eobacteria; Betaproteobacteria; Burkholderiales;
ae; Ralstonia.
N.A.
379; PubMed=11823852;
Genin S., Artiguenave F., Gouzy J., Mangenot S.,
```

```
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein. Extracellular
CC       matrix.
CC   -!- TISSUE SPECIFICITY: Oocytes.
CC   -!- DEVELOPMENTAL STAGE: Expressed during the 2-week growth phase
CC       oogenesis, prior to ovulation.
CC   -!- PTM: Sulfated glycoprotein with O-linked oligosaccharides.
CC   -!- SIMILARITY: Contains 1 ZP domain.
CC
CC   This SWISS-PROT entry is copyright. It is produced through a col
CC   between the Swiss Institute of Bioinformatics and the EMBL out
CC   the European Bioinformatics Institute. There are no restrictio
CC   use by non-profit institutions as long as its content is ir
CC   modified and this statement is not removed. Usage by and for c
CC   entities requires a license agreement (See http://www.isb-sib.ch/
CC   or send an email to license@isb-sib.ch).
CC
DR   EMBL; M20026; AAB18629.1; -.
DR   EMBL; X14376; CAA32550.1; -.
DR   PIR; A30334; A30334.
DR   MGD; MGI:99215; Zp3.
DR   InterPro; IPR001507; Endoglin/CD105.
DR   Pfam; PF00100; zona_pellucida; 1.
DR   PRINTS; PR00023; ZPELLUCIDA.
DR   SMART; SM00241; ZP; 1.
DR   PROSITE; PS00682; ZP_DOMAIN; 1.
KW   Glycoprotein; Signal; Sulfation; Sperm; Receptor; Transmembrane;
KW   Extracellular matrix.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23    424       ZONA PELLUCIDA SPERM-BINDING PROTEIN
FT   DOMAIN       23    387       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    388    408       POTENTIAL.
FT   DOMAIN      409    424       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       45    308       ZP.
FT   CARBOHYD    146    146       N-LINKED (GLCNAC...)   (POTENTIAL).
FT   CARBOHYD    273    273       N-LINKED (GLCNAC...)   (POTENTIAL).
FT   CARBOHYD    304    304       N-LINKED (GLCNAC...)   (POTENTIAL).
FT   CARBOHYD    327    327       N-LINKED (GLCNAC...)   (POTENTIAL).
FT   CARBOHYD    330    330       N-LINKED (GLCNAC...)   (POTENTIAL).
SQ   SEQUENCE    424 AA;  46303 MW;  9089903FBD268365 CRC64;

Query Match             2.5%;  Score 7;  DB 1;  Length 424;
Best Local Similarity  100.0%;  Pred. No. 96;
Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;

Qy          60 ALGLGLA 66
               |||||||
Db         389 ALGLGLA 395

RESULT 72
SYH_CHLMU
ID   SYH_CHLMU      STANDARD;      PRT;    428 AA.
AC   Q9PUJ9;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Histidyl-tRNA synthetase (EC 6.1.1.21) (Histidine--tRNA ligase)
DE   (HisRS).
GN   HISS OR TC0830.
OS   Chlamydia muridarum.
OC   Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=83560;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MoPn / Nigg;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
```

```
S57103.
41916; -.
YJR084W.
0; C:signalosome complex; IDA.
4; P:adaptation to pheromone during conjugation ...; IMP.
000717; PCI.
; PCI; 1.
protein.
3 AA;  49482 MW;  750CDA631916A621 CRC64;

arity  100.0%;  Pred. No. 95;  DB 1;  Length 423;
         Score 7;
onservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

LAL 61
|||
LAL 272

STANDARD;      PRT;    424 AA.

Rel. 11, Created)
Rel. 35, Last sequence update)
Rel. 39, Last annotation update)
a sperm-binding protein 3 precursor (Zona pellucida
ZP3) (Sperm receptor) (Zona pellucida protein C).

(Mouse)
tazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
heria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
090;

N.A.
926; PubMed=3378665;
., Chamberlin M.E., Baur A.W., Sobieski D.A., Dean J.;
alysis of cDNA coding for ZP3, a sperm binding protein
zona pellucida.";
7:287-295(1988).

87.

V-1996) to the EMBL/GenBank/DDBJ databases.

N.A.
TISSUE=liver;
048; PubMed=2541416;
Wassarman P.M.;
equence of the gene encoding zona pellucida glycoprotein
use sperm receptor.";
Res. 17:2861-2863(1989).

N.A.
451; PubMed=2842770;
Roller R.J., Fimiani C.M., Wassarman D.A.,
.;
cture of the mouse sperm receptor polypeptide determined
oning.";
cad. Sci. U.S.A. 85:6409-6413(1988).

9-63; 197-204; 219-233 AND 261-275.

0795; PubMed=1330788;
Wassarman P.M.;
on of a region of mouse zona pellucida glycoprotein mZP3
es sperm receptor activity.";
4:309-317(1992).
Functions as a sperm-receptor. It is responsible for
esion to the zona pellucida, and may contribute to the
specificity of the insemination.
ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN
? AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
```

Res. 28:1397-1406 (2000).
ACTIVITY: ATP + L-histidine + tRNA(His) = AMP +
e + L-histidyl-tRNA(His).
omodimer (By similarity).
R LOCATION: Cytoplasmic.
: Belongs to class-II aminoacyl-tRNA synthetase family.

T entry is copyright. It is produced through a collaboration
wiss Institute of Bioinformatics and the EMBL outstation -
ioinformatics Institute. There are no restrictions on its
rofit institutions as long as its content is in no way
his statement is not removed. Usage by and for commercial
res a license agreement (See http://www.isb-sib.ch/announce/
il to license@isb-sib.ch).

; AAF39630.1; -.
81660.
1QE0.

-; -; 1.
04154; HGTP_anticodon.
04516; HisS.
02314; tRNA-synt_2b.
06195; tRNA_ligase_II.
HGTP_anticodon; 1.
tRNA-synt_2b; 1.
00442; hisS; 1.
62; AA_TRNA_LIGASE_II; 1.
synthetase; Protein biosynthesis; Ligase; ATP-binding;
ome.
AA; 48939 MW; 9CF859ED0B689DDF CRC64;

rity 2.5%; Score 7; DB 1; Length 428;
100.0%; Pred. No. 96;
nservative 0; Mismatches 0; Indels 0; Gaps 0;

AR 139
|\
AR 181

STANDARD; PRT; 455 AA.

el. 12, Created)
el. 12, Last sequence update)
el. 31, Last annotation update)
imidine photolyase (EC 4.1.99.3) (DNA photolyase)
ting enzyme).

riseus.
nobacteria; Actinobacteridae; Actinomycetales;
e; Streptomycetaceae; Streptomyces.
1;

N.A.
14; PubMed=2501760;
Takao M., Oikawa A., Yasui A.;
racterization of a gene encoding a photolyase from
riseus.";
Res. 17:4731-4744(1989).
This enzyme catalyzes the light-dependent monomerization
M) of cyclobutyl pyrimidine dimers (in cis-syn
ion), which are formed between adjacent bases on the
strand, upon exposure to ultraviolet radiation.
ACTIVITY: Cyclobutadipyrimidine (in DNA) = 2 pyrimidine
(in DNA).
Contains 2 chromophores: a reduced flavin (FADH2) and an
-hydroxy-5-deazaflavin (F420). Both chromophores are
on-covalent interactions.
OUS: THIS PROTEIN BELONGS TO THE "LONG WAVELENGTH-TYPE
IS" WITH AN ABSORPTION MAXIMUM AT ABOUT 440 NM.

-!- SIMILARITY: Belongs to the DNA photolyase class-1 family.

CC
CC This SWISS-PROT entry is copyright. It is produced through a colla
CC between the Swiss Institute of Bioinformatics and the EMBL outs
CC the European Bioinformatics Institute. There are no restriction:
CC use by non-profit institutions as long as its content is in
CC modified and this statement is not removed. Usage by and for co
CC entities requires a license agreement (See http://www.isb-sib.ch/a
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X15060; CAA33161.1; -.
DR PIR; S05573; S05573.
DR HSSP; P00914; 1DNP.
DR InterPro; IPR002081; DNA_photolyase_1.
DR InterPro; IPR006050; DNA_photolyase_N.
DR InterPro; IPR005101; FAD_binding_7.
DR InterPro; IPR006051; FAD_binding_N.
DR Pfam; PF00875; DNA_photolyase; 1.
DR Pfam; PF03441; FAD_binding_7; 1.
DR PRINTS; PR00147; DNAPHOTLYASE.
DR ProDom; PD004390; FAD_binding_N; 1.
DR PROSITE; PS00394; DNA_PHOTOLYASES_1_1; 1.
DR PROSITE; PS00691; DNA_PHOTOLYASES_1_2; 1.
KW Lyase; Chromophore; Flavoprotein; FAD; DNA repair; DNA-binding.
FT DNA_BIND 300 319 H-T-H MOTIF (POTENTIAL).
SQ SEQUENCE 455 AA; 50436 MW; FEA88763DA45931A CRC64;

Query Match 2.5%; Score 7; DB 1; Length 455;
Best Local Similarity 100.0%; Pred. No. 1e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Ga

Qy 214 LDLLVDG 220
     |||||||
Db 356 LDLLVDG 362

RESULT 74
PUCC_RHOCA
ID PUCC_RHOCA STANDARD; PRT; 461 AA.
AC P23462;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein pucC.
GN PUCC.
OS Rhodobacter capsulatus (Rhodopseudomonas capsulata).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Rhodobacter.
OX NCBI_TaxID=1061;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89359127; PubMed=2549005;
RA Tichy H.V., Oberle B., Stiehle H., Schiltz E., Drews G.;
RT "Genes downstream from pucB and pucA are essential for formation o
RT the B800-850 complex of Rhodobacter capsulatus.";
RL J. Bacteriol. 171:4914-4922(1989).
RN [2]
RP REVISION, AND FUNCTION.
RX MEDLINE=92007739; PubMed=1717257;
RA Tichy H.V., Albien K.-U., Gad'On N., Drews G.;
RT "Analysis of the Rhodobacter capsulatus puc operon: the pucC gene
RT plays a central role in the regulation of LHII (B800-850 complex)
RT expression.";
RL EMBO J. 10:2949-2955(1991).
RN [3]
RP TOPOLOGY.
RX MEDLINE=96326322; PubMed=8759841;
RA LeBlanc H.N., Beatty J.T.;
RT "Topological analysis of the Rhodobacter capsulatus PucC protein a
RT effects of C-terminal deletions on light-harvesting complex II.";
RL J. Bacteriol. 178:4801-4806(1996).
CC -!- FUNCTION: PUCC IS REQUIRED FOR HIGH-LEVEL TRANSCRIPTION OF TH
CC PUC OPERON.

AR LOCATION: Integral membrane protein.

OT entry is copyright. It is produced through a collaboration
Swiss Institute of Bioinformatics and the EMBL outstation -
Bioinformatics Institute. There are no restrictions on its
profit institutions as long as its content is in no way
this statement is not removed. Usage by and for commercial
ires a license agreement (See http://www.isb-sib.ch/announce/
ail to license@isb-sib.ch).

AAA26163.1; ALT_INIT.
C33958.
004896; PUCC.
; PUCC; 1.
ex; Transmembrane.

1	36	CYTOPLASMIC (PROBABLE).
37	56	PROBABLE.
57	62	PERIPLASMIC (PROBABLE).
63	83	PROBABLE.
84	109	CYTOPLASMIC (PROBABLE).
10	129	PROBABLE.
30	142	PERIPLASMIC (PROBABLE).
43	163	PROBABLE.
64	182	CYTOPLASMIC (PROBABLE).
83	202	PROBABLE.
03	209	PERIPLASMIC (PROBABLE).
10	228	PROBABLE.
29	261	CYTOPLASMIC (PROBABLE).
62	281	PROBABLE.
82	302	PERIPLASMIC (PROBABLE).
03	319	PROBABLE.
20	338	CYTOPLASMIC (PROBABLE).
39	355	PROBABLE.
56	358	PERIPLASMIC (PROBABLE).
59	376	PROBABLE.
77	394	CYTOPLASMIC (PROBABLE).
95	415	PROBABLE.
16	436	PERIPLASMIC (PROBABLE).
37	456	PROBABLE.
57	461	CYTOPLASMIC (PROBABLE).

1 AA; 48392 MW; BFC7A8A0C549875A CRC64;

arity   100.0%; Pred. No. 1e+02; Indels   0; Gaps   0;
onservative   0; Mismatches   0;

EDQ 101
|||
EDQ 176

STANDARD;   PRT;   461 AA.

(Rel. 35, Created)
(Rel. 35, Last sequence update)
(Rel. 41, Last annotation update)
protein HI0608.

nfluenzae.
oteobacteria; Gammaproteobacteria; Pasteurellales;
ae; Haemophilus.
27;

M.N.A.
KW20 / ATCC 51907;
0630; PubMed=7542800;
R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
, Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,

RA   Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.(
RA   Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA   Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA   Venter J.C.;
RT   "Whole-genome random sequencing and assembly of Haemophilus infl
RT   Rd.";
RL   Science 269:496-512(1995).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- SIMILARITY: BELONGS TO THE SLC13A FAMILY OF TRANSPORTERS.
CC   NADC SUBFAMILY.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a coll
CC   between the Swiss Institute of Bioinformatics and the EMBL out
CC   the European Bioinformatics Institute. There are no restrictior
CC   use by non-profit institutions as long as its content is li
CC   modified and this statement is not removed. Usage by and for (
CC   entities requires a license agreement (See http://www.isb-sib.ch/
CC   or send an email to license@isb-sib.ch).
DR   EMBL; U32743; AAC22267.1; -.
DR   PIR; I64080; I64080.
DR   TIGR; HI0608; -.
DR   InterPro; IPR001898; Na/sul_symport.
DR   Pfam; PF00939; Na_sulph_symp; 1.
DR   TIGRFAMs; TIGR00785; dass; 1.
DR   PROSITE; PS01271; NA_SULFATE; 1.
KW   Hypothetical protein; Transmembrane; Transport; Complete proteome
FT   TRANSMEM   13   33   POTENTIAL.
FT   TRANSMEM   54   74   POTENTIAL.
FT   TRANSMEM   81   101   POTENTIAL.
FT   TRANSMEM   120   140   POTENTIAL.
FT   TRANSMEM   170   190   POTENTIAL.
FT   TRANSMEM   211   231   POTENTIAL.
FT   TRANSMEM   256   276   POTENTIAL.
FT   TRANSMEM   286   306   POTENTIAL.
FT   TRANSMEM   314   334   POTENTIAL.
FT   TRANSMEM   349   369   POTENTIAL.
FT   TRANSMEM   377   397   POTENTIAL.
FT   TRANSMEM   399   419   POTENTIAL.
FT   TRANSMEM   439   459   POTENTIAL.
SQ   SEQUENCE   461 AA; 49761 MW; B5E6F965B38EF06 CRC64;

Query Match   2.5%; Score 7; DB 1; Length 461;
Best Local Similarity   100.0%; Pred. No. 1e+02; Indels   0;
Matches   7; Conservative   0; Mismatches   0; Indels   0; (

Qy   53 TALLVPL 59
               |||||||
Db   60 TALLVPL 66

Search completed: April 7, 2004, 17:57:56
Job time : 21 secs

GenCore version 5.1.6
opyright (c) 1993 - 2004 Compugen Ltd.

n search, using sw model

il 7, 2004, 17:54:13 ; Search time 45 Seconds
                        (without alignments)
                        1991.270 Million cell updates/sec

-09-245-198A-4

SLLDFEISARRLPLPRSLG.........PWAHLKAAPFLTYFGLFQVH 284

GO

op 60.0 , Gapext 60.0

7041 seqs, 315518202 residues

s satisfying chosen parameters: 1017041

gth: 0
Max gth: 2000000000

sting first 100 summaries

TREMBL_25:*
sp_archea:*
sp_bacteria:*
sp_fungi:*
sp_human:*
sp_invertebrate:*
sp_mammal:*
sp_mhc:*
sp_organelle:*
sp_phage:*
sp_plant:*
sp_rodent:*
sp_virus:*
sp_vertebrate:*
sp_unclassified:*
sp_rvirus:*
sp_bacteriap:*
sp_archeap:*

the number of results predicted by chance to have a
r than or equal to the score of the result being printed,
ed by analysis of the total score distribution.

SUMMARIES

ery				
ch	Length	DB	ID	Description
3.5	330	4	Q8IZK7	Q8izk7 homo sapien
1.3	410	11	Q8BXS2	Q8bxs2 mus musculu
1.2	438	16	Q7VVB7	Q7vvb7 bordetella
4.2	470	16	Q7W7F2	Q7w7f2 bordetella
3.5	111	16	Q8X4J8	Q8x4j8 escherichia
3.5	111	16	Q8FFL8	Q8ffl8 escherichia
3.5	111	16	Q7UC61	Q7uc61 shigella fl
3.2	142	16	Q98IJ0	Q98ij0 rhizobium l
3.2	749	16	Q7V5I1	Q7v5i1 prochloroco
3.2	766	16	Q8PPZ8	Q8ppz8 xanthomonas
3.2	1208	16	Q7ULK4	Q7ulk4 rhodopirell
2.8	143	17	Q9HST7	Q9hst7 halobacteri
2.8	151	10	Q9SDI1	Q9sdi1 oryza sativ
2.8	154	2	Q848K4	Q848k4 gamma-prote
2.8	190	16	Q8E569	Q8e569 streptococc
2.8	190	16	Q8DZK5	Q8dzk5 streptococc

#		Score	Length	DB	ID	Accession/Organism
17	8	2.8	193	10	Q8L4K2	Q8l4k2 ory
18	8	2.8	197	16	Q9WYU0	Q9wyu0 the
19	8	2.8	211	16	Q9VV70	Q9vv70 dros
20	8	2.8	211	5	Q8SXH4	Q8sxh4 dros
21	8	2.8	220	16	Q8FV59	Q8fv59 bru
22	8	2.8	232	16	Q89GW9	Q89gw9 bra
23	8	2.8	276	10	Q84MB7	Q84mb7 ara
24	8	2.8	278	16	Q92G99	Q9zg99 pse
25	8	2.8	278	16	Q82J43	Q82j43 str
26	8	2.8	279	11	Q9D378	Q9d378 mus
27	8	2.8	279	11	Q9CPR8	Q9cpr8 mus
28	8	2.8	306	5	Q9VNP0	Q9vnp0 dros
29	8	2.8	317	6	Q864L1	Q864l1 pan
30	8	2.8	317	6	Q864L0	Q864l0 pong
31	8	2.8	317	6	Q864J8	Q864j8 maca
32	8	2.8	317	6	Q864J7	Q864j7 maca
33	8	2.8	317	6	Q864J5	Q864j5 maca
34	8	2.8	317	6	Q864G8	Q864g8 aotu
35	8	2.8	317	6	Q864G7	Q864g7 atel
36	8	2.8	317	6	Q864G6	Q864g6 alou
37	8	2.8	317	6	Q864G4	Q864g4 alou
38	8	2.8	317	6	Q864G3	Q864g3 alou
39	8	2.8	317	6	Q864G2	Q864g2 alou
40	8	2.8	317	6	Q864G1	Q864g1 alou
41	8	2.8	317	6	Q864F8	Q864f8 vare
42	8	2.8	317	6	Q864F7	Q864f7 vare
43	8	2.8	317	6	Q864F6	Q864f6 eule
44	8	2.8	317	6	Q864F5	Q864f5 hapa
45	8	2.8	317	6	Q864F4	Q864f4 lemu
46	8	2.8	339	17	O58554	O58554 pyr
47	8	2.8	342	8	Q99908	Q99908 homo
48	8	2.8	343	4	Q9BPV2	Q9bpv2 homo
49	8	2.8	370	17	Q8TUU8	Q8tuu8 met
50	8	2.8	372	16	O53860	O53860 myc
51	8	2.8	372	16	Q7U145	Q7u145 myc
52	8	2.8	375	1	Q877C5	Q877c5 meth
53	8	2.8	377	16	Q9KYT3	Q9kyt3 str
54	8	2.8	387	10	Q9SHD8	Q9shd8 ara
55	8	2.8	397	16	Q8U8Z0	Q8u8z0 agr
56	8	2.8	431	16	Q7U9H4	Q7u9h4 syn
57	8	2.8	435	16	Q8XS70	Q8xs70 ral
58	8	2.8	465	16	Q8ZCV8	Q8zcv8 yer
59	8	2.8	471	16	Q9RRI8	Q9rri8 dei
60	8	2.8	472	4	Q96N66	Q96n66 homo
61	8	2.8	473	11	Q8R1P9	Q8r1p9 mus
62	8	2.8	473	11	Q9CY76	Q9cy76 mus
63	8	2.8	473	11	Q8CHK3	Q8chk3 mus
64	8	2.8	496	10	Q8GVZ8	Q8gvz8 ory
65	8	2.8	522	10	Q9LIW0	Q9liw0 ory
66	8	2.8	535	16	Q7WZ71	Q7wz71 nono
67	8	2.8	537	16	Q8PMH8	Q8pmh8 xan
68	8	2.8	564	12	Q9YTU9	Q9ytu9 inf
69	8	2.8	568	12	Q8JN92	Q8jn92 inf
70	8	2.8	568	12	Q8QPL0	Q8qpl0 inf
71	8	2.8	620	16	Q8PAS2	Q8pas2 xan
72	8	2.8	656	5	Q9N8H2	Q9n8h2 tryp
73	8	2.8	1523	5	Q9N8U8	Q9n8u8 tryp
74	7	2.5	32	6	Q9BEF6	Q9bef6 capr
75	7	2.5	35	4	Q9ZG35	Q9zg35 chla
76	7	2.5	54	5	Q16193	Q16193 homo
77	7	2.5	54	16	Q7U6G4	Q7u6g4 syn
78	7	2.5	67	2	Q8RQS4	Q8rqs4 serr
79	7	2.5	69	10	Q8H142	Q8h142 ara
80	7	2.5	86	16	Q9A4G4	Q9a4g4 cau
81	7	2.5	94	4	Q9P074	Q9p074 homo
82	7	2.5	97	11	Q82TB6	Q82tb6 nit
83	7	2.5	98	10	Q7X682	Q7x682 ory
84	7	2.5	99	2	P97199	P97199 esch
85	7	2.5	100	12	Q8WQR8	Q8wqr8 eupr
86	7	2.5	100	15	Q8QRH6	Q8qrh6 hep
87	7	2.5	101	7	Q8UTL1	Q8utl1 hum
88	7	2.5	104	9	Q8DBE3	Q8dbe3 vib
89	7	2.5	108	9	Q38214	Q38214 bact

```
2.5  115  16  Q9HY60    Q9hy60 pseudomonas
2.5  118   2  Q939F7    Q939f7 pseudomonas
2.5  118  16  Q928G5    Q928g5 listeria in
2.5  119  10  Q94LX6    Q941x6 euphorbia in
2.5  121  10  Q8W0M1    Q8w0m1 oryza sativ
2.5  122  10  Q84PV3    Q84pv3 oryza sativ
2.5  122  16  Q8EFG2    Q8efg2 shewanella
2.5  123  16  Q8UGH2    Q8ugh2 agrobacteri
2.5  123  16  Q7U8Z3    Q7u8z3 synechococc
2.5  124  10  Q7XRK7    Q7xrk7 oryza sativ
2.5  125   5  Q9U1P6    Q9u1p6 caenorhabdi
```

ALIGNMENTS

PRELIMINARY;   PRT;   330 AA.

TrEMBLrel. 23, Created)
TrEMBLrel. 23, Last sequence update)
TrEMBLrel. 25, Last annotation update)

(Human).
...tazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
...heria; Primates; Catarrhini; Hominidae; Homo.
...06;
...N.A.
...924; PubMed=12411489;
... B., Medema J.P., Lopez-Fraga M., Lozano J.C.,
...M., Picard A., Martinez-A C., Garcia-Sanz J.A.,
...s hybrid mRNA encodes TWE-PRIL, a functional cell surface
...usion protein.";
...11-5720(2002).
...1; AAL90443.1; -.
...0; C:membrane; IEA.
...4; F:tumor necrosis factor receptor binding; IEA.
...5; P:immune response; IEA.
...006052; TNF_family.
...008983; TNF_like.
...; TNF; 1.
...7; TNF; 1.
...1251; TNF_1; 1.
...1049; TNF_2; 2.
...0 AA;   36588 MW;   FC6F3BCA29C029AE CRC64;

...arity   58.5%;   Score 166;   DB 4;   Length 330;
...100.0%;   Pred. No. 1.3e-155;
...onservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
RSQRRGRRGEPGTALLVPLALGLGHLALACLGLLLAVVSLGSRASLSAQEPAQEBL 95
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
RSQRRGRRGEPGTALLVPLALGLGHLALACLGLLLAVVSLGSRASLSAQEPAQEBL 60

DQDPSELNPQTEESQDPAPPLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQD 155
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DQDPSELNPQTEESQDPAPPLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQD 120

AGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQ 201
|||||||||||||||||||||||||||||||||||||||||||
AGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQ 166
```

PRELIMINARY;   PRT;   410 AA.

(TrEMBLrel. 23, Created)
(TrEMBLrel. 23, Last sequence update)
(TrEMBLrel. 25, Last annotation update)

DE   Tumor necrosis factor.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Retina;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium;
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotati
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK044387; BAC31897.1; -.
DR   PIR; PT0714; PT0714.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005164; F:tumor necrosis factor receptor binding; IEA.
DR   GO; GO:0006955; P:immune response; IEA.
DR   InterPro; IPR006052; TNF_family.
DR   InterPro; IPR008983; TNF_like.
DR   SMART; SM00207; TNF; 2.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 2.
SQ   SEQUENCE   410 AA;   45881 MW;   590A4B74C33FB8D4 CRC64;

Query Match   11.3%;   Score 32;   DB 11;   Length 410;
Best Local Similarity   100.0%;   Pred. No. 8.9e-23;
Matches   32;   Conservative   0;   Mismatches   0;   Indels   0;   C

```
Qy   139 RRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEE 170
             ||||||||||||||||||||||||||||||||
Db   105 RRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEE 136
```

RESULT 3
Q7VVB7
ID   Q7VVB7   PRELIMINARY;   PRT;   438 AA.
AC   Q7VVB7;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Putative chloride-channel protein.
GN   BP2760.
OS   Bordetella pertussis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Alcaligenaceae; Bordetella.
OX   NCBI_TaxID=520;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RX   MEDLINE=22827954; PubMed=12910271;
RA   Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,
RA   Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall
RA   Cerdeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.
RA   Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I
RA   Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA   Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels
RA   Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price
RA   Rabbinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA   Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Steve
RA   Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT   "Comparative analysis of the genome sequences of Bordetella pertu
RT   Bordetella parapertussis and Bordetella bronchiseptica.";
RL   Nat. Genet. 35:32-40(2003).
DR   EMBL; BX640419; CAE43035.1; -.
KW   Complete proteome.
SQ   SEQUENCE   438 AA;   45402 MW;   AD51A65B595D99D8 CRC64;

Query Match   4.2%;   Score 12;   DB 16;   Length 438;
Best Local Similarity   100.0%;   Pred. No. 0.0062;
Matches   12;   Conservative   0;   Mismatches   0;   Indels   0;   C

```
DR   EMBL; AE005458; AAG57389.1; -.
DR   PIR; A85866; A85866.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR000620; DUF6.
DR   Pfam; PF00892; DUF6; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   111 AA;  12165 MW;  7CEFC93D786CD759 CRC64;
```

```
Query Match             3.5%;  Score 10;  DB 16;  Length 111;
Best Local Similarity  100.0%;  Pred. No. 0.17;
Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  G
```

```
Qy         63 LGLALACLGL 72
               ||||||||||
Db         40 LGLALACLGL 49
```

```
RESULT 6
Q8FFL8
ID   Q8FFL8           PRELIMINARY;       PRT;   111 AA.
AC   Q8FFL8;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Conserved hypothetical protein.
GN   C2800.
OS   Escherichia coli O6.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=217992;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX   MEDLINE=22388234; PubMed=12471157;
RA   Welch R.A., Burland V., Plunkett G.  III, Redford P., Roesch P.,
RA   Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud
RA   Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA   Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT   "Extensive mosaic structure revealed by the complete genome seque
RT   of uropathogenic Escherichia coli.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
DR   EMBL; AE016763; AAN81254.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   111 AA;  12196 MW;  C0A977B6F77AB87 CRC64;
```

```
Query Match             3.5%;  Score 10;  DB 16;  Length 111;
Best Local Similarity  100.0%;  Pred. No. 0.17;
Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  G
```

```
Qy         63 LGLALACLGL 72
               ||||||||||
Db         40 LGLALACLGL 49
```

```
RESULT 7
Q7UC61
ID   Q7UC61           PRELIMINARY;       PRT;   111 AA.
AC   Q7UC61;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Sucrose-6 phosphate hydrolase.
GN   S2567.
OS   Shigella flexneri.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Shigella.
OX   NCBI_TaxID=623;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX   MEDLINE=22590274; PubMed=12704152;
RA   Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA   Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
```

Left column (rotated fragments):

```
LACLGLL 73
   ||||||
LACLGLL 285
```

```
PRELIMINARY;       PRT;   470 AA.

(TrEMBLrel. 25, Created)
(TrEMBLrel. 25, Last sequence update)
(TrEMBLrel. 25, Last annotation update)
ide-channel protein.

apertussis.
eobacteria; Betaproteobacteria; Burkholderiales;
; Bordetella.
;

N.A.
/ ATCC BAA-587;
54; PubMed=12910271;
ebaihia M., Preston A., Murphy L.D., Thomson N.,
olden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
ja A.M., Temple L., James K., Harris B., Quail M.A.,
kin R., Baker S., Basham D., Bason N., Cherevach I.,
T., Collins M., Cronin A., Davis P., Doggett J.,
oble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
ule S., Norberczak H., O'Neil S., Ormond D., Price C.,
E., Rutter S., Sanders M., Saunders D., Seeger K.,
onds M., Skelton J., Squares R., Squares S., Stevens K.,
ehead S., Barrell B.G., Maskell D.J.;
analysis of the genome sequences of Bordetella pertussis,
apertussis and Bordetella bronchiseptica.";
5:32-40(2003).
; CAE37862.1; -.
eome.
 AA;  49233 MW;  5B92DD05E920BD5F CRC64;
```

```
arity             4.2%;  Score 12;  DB 16;  Length 470;
                 100.0%;  Pred. No. 0.0066;
onservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
LACLGLL 73
   ||||||
LACLGLL 317
```

```
PRELIMINARY;       PRT;   111 AA.

(TrEMBLrel. 20, Created)
(TrEMBLrel. 20, Last sequence update)
(TrEMBLrel. 24, Last annotation update)
protein z3516.

oli O157:H7.
eobacteria; Gammaproteobacteria; Enterobacteriales;
aceae; Escherichia.
334;

N.A.
7 / EDL933 / ATCC 700927;
935; PubMed=11206551;
lunkett G. III, Burland V., Mau B., Glasner J.D.,
yhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
ckett J., Klink S., Boutin A., Shao Y., Miller L.,
Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
nantharaman T.S., Lin J., Yen G., Schwartz D.C.,
lattner F.R.;
nce of enterohaemorrhagic Escherichia coli O157:H7.";
9-533(2001).
```

RA ...N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA ..., Blattner F.R.;
RT ...nome sequence and comparative genomics of Shigella
RT ...otype 2a strain 2457T.";
RL ...1.71:2775-2786(2003).
DR ...36; AAP17670.1; -.

11 AA; 12224 MW; 7CFA06CC46A32672 CRC64;

larity 3.5%; Score 10; DB 16; Length 111;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   ALACLGL 72
          |||||||
Db   ALACLGL 49

PRELIMINARY; PRT; 142 AA.

(TrEMBLrel. 18, Created)
(TrEMBLrel. 18, Last sequence update)
(TrEMBLrel. 18, Last annotation update)
...sport protein.
...ti (Mesorhizobium loti).
...oteobacteria; Alphaproteobacteria; Rhizobiales;
...iaceae; Mesorhizobium.
...81;
M N.A.
03099;
2930; PubMed=11214968;
...akamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
...Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
...Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
..., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
...Yamada M., Tabata S.;
...nome structure of the nitrogen-fixing symbiotic bacterium
...m loti.";
...31-338(2000).
...99; BAB49526.1; -.
...teome.
42 AA; 14884 MW; 0DCA7842CB5A5E6F CRC64;

larity 3.2%; Score 9; DB 16; Length 142;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   LVPLAL 61
          ||||||
Db   LVPLAL 130

PRELIMINARY; PRT; 749 AA.

(TrEMBLrel. 25, Created)
(TrEMBLrel. 25, Last sequence update)
(TrEMBLrel. 25, Last annotation update)
I PsaB protein.
cus marinus (strain MIT 9313).
anobacteria; Prochlorophytes; Prochlorococcaceae;
cus.
4547;
M N.A.
5698; PubMed=12917642;
rimer F.W., Lamerdin J., Malfatti S., Chain P.,

RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
RA Johnson Z.I., Steglich C., Land M., Lindell D., Post A.F., Regala W., Shah M.,
RA Shaw S.L., Zinser E.R., Chisholm S.W.;
RT "Genome divergence in two Prochlorococcus ecotypes reflects ocea
RT niche differentiation.";
RL Nature 424:1042-1047(2003).
DR EMBL; BX572100; CAE21944.1; -.
KW Photosystem I; Complete proteome.
SQ SEQUENCE 749 AA; 83231 MW; B1D49645F1C790C CRC64;

Query Match 3.2%; Score 9; DB 16; Length 749;
Best Local Similarity 100.0%; Pred. No. 9.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0;

Qy 63 LGlALACLG 71
          |||||||||
Db 347 LGlALACLG 355

RESULT 10
Q8PPZ8
ID Q8PPZ8   PRELIMINARY;   PRT;   766 AA.
AC Q8PPZ8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE C-type cytochrome biogenesis protein (Copper tolerance).
GN DSBD OR XAC0534.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22021145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with dif
RT host specificities.";
RL Nature 417:459-463(2002).
DR EMBL; AE011680; AAM35423.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0017004; P:cytochrome biogenesis; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR003834; Cytoch_TM.
DR InterPro; IPR006662; Thiored_
DR InterPro; IPR006663; Thioredox_dom2.
DR Pfam; PF02683; DsbD; 1.
DR PROSITE; PS00194; THIOREDOXIN; 1.
KW Complete proteome.
SQ SEQUENCE 766 AA; 81014 MW; 3A1955A07DB8A9CA CRC64;

Query Match 3.2%; Score 9; DB 16; Length 766;
Best Local Similarity 100.0%; Pred. No. 9.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0;

Qy 59 LALGLGLAL 67

```
Query Match          2.8%;     Score 8;  DB 17;  Length 143;
Best Local Similarity  100.0%;  Pred. No. 21;
Matches    8;  Conservative   0;  Mismatches   0;  Indels   0;  G

Qy    61  LGLGLALA  68
          ||||||||
Db    55  LGLGLALA  62

RESULT 13
Q9SDI1
ID   Q9SDI1        PRELIMINARY;      PRT;    151 AA.
AC   Q9SDI1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein (OSJNBA0036E02.6 protein) (B1085F09.2
DE   protein).
GN   B1085F09.2.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT   clone:P0003H10.";
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
RT   clone:OSJNBA0036E02.";
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
RT   clone:B1085F09.";
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP000815; BAA87834.1; -.
DR   EMBL; AP002862; BAB17732.1; -.
DR   EMBL; AP003103; BAB44106.1; -.
DR   Gramene; Q9SDI1; -.
SQ   SEQUENCE   151 AA;  16632 MW;  EC68451ECA2BD71D CRC64;

Query Match          2.8%;     Score 8;  DB 10;  Length 151;
Best Local Similarity  100.0%;  Pred. No. 22;
Matches    8;  Conservative   0;  Mismatches   0;  Indels   0;  G

Qy    43  RRRGRRGE  50
          ||||||||
Db   131  RRRGRRGE  138

RESULT 14
Q848K4
ID   Q848K4        PRELIMINARY;      PRT;    154 AA.
AC   Q848K4;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein (Fragment).
OS   Gamma-proteobacterium Hot 75m4.
OG   Plasmid pAK106.
OC   Bacteria; environmental samples.
OX   NCBI_TaxID=77133;
RN   [1]
```

```
||||
GLAL 542

ELIMINARY;      PRT;    1208 AA.

rEMBLrel. 25, Created)
rEMBLrel. 25, Last sequence update)
rEMBLrel. 25, Last annotation update)
 multi-functional protein.

 baltica.
nctomycetes; Planctomycetacia; Planctomycetales;
eae; Pirellula.
;

N.A.

13; PubMed=12835416;
, Kube M., Bauer M., Teeling H., Lombardot T.,
le D., Beck A., Borzym K., Heitmann K., Rabus R.,
Amann R., Reinhardt R.;
ome sequence of the marine planctomycete Pirellula sp.

cad. Sci. U.S.A. 100:8298-8303(2003).
); CAD76265.1; -.
eome.
08 AA;  132047 MW;  0FFE225741021E8C CRC64;

arity  100.0%;   Pred. No. 15;
onservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

AQEPA 91
||||
AQEPA 42

ELIMINARY;      PRT;    143 AA.

rEMBLrel. 16, Created)
rEMBLrel. 16, Last sequence update)
rEMBLrel. 24, Last annotation update)

sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
archaeota; Halobacteria; Halobacteriales;
eae; Halobacterium.
091;

N.A.
483; PubMed=11016950;
edy S.P., Mahairas G.G., Berquist B., Pan M.,
Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
, Keller K., Cruz R., Danson M.J., Hough D.W.,
, Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
tas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
owe T.M., Liang P., Riley M., Hood L., DaSarma S.;
nce of Halobacterium species NRC-1.";
cad. Sci. U.S.A. 97:12176-12181(2000).
6; AAG18715.1; -.
G84168.
006976; VanZ.
; VanZ; 1.
eome.
3 AA;  15648 MW;  45466E6328EF3468 CRC64;
```

```
1 N.A.
7661; PubMed=12620823;
Waschkowitz T., Bowien S., Henne A., Daniel R.;
and Screening of Metagenomic Libraries Derived from
ltures: Generation of a Gene Bank for Genes Conferring
reductase Activity on Escherichia coli.";
. Microbiol. 69:1408-1416(2003).
30; AAO59972.1; -.
21; C:extrachromosomal DNA; IEA.
protein; Plasmid.
1     1                                               54 AA;  16234 MW;  3AE8A072D5B7E137 CRC64;

larity   2.8%;  Score 8;  DB 2;  Length 154;
Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
RGRRG 49
|||||
RGRRG 119
```

```
PRELIMINARY;      PRT;    190 AA.

(TrEMBLrel. 23, Created)
(TrEMBLrel. 23, Last sequence update)
(TrEMBLrel. 24, Last annotation update)
protein.

s agalactiae (serotype III).
rmicutes; Lactobacillales; Streptococcaceae;
s.
16495;
```

```
M N.A.
6 / Serotype III;
2508; PubMed=12354221;
usniok C., Buchrieser C., Chevalier F., Frangeul L.,
ouine M., Couve E., Lalioui L., Poyart C., Trieu-Cuot P.,
ence of Streptococcus agalactiae, a pathogen causing
natal disease.";
ol. 45:1499-1513(2002).
49; CAD46822.1; -.
s1163; -.
R008172; Adenylate_cyc.
8; CYTH; 1.
protein; Complete proteome.
90 AA;  22178 MW;  AB2AD33C2CB6FBF3 CRC64;

larity   2.8%;  Score 8;  DB 16;  Length 190;
Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
IRTLP 265
|||||
IRTLP 57
```

```
PRELIMINARY;      PRT;    190 AA.

(TrEMBLrel. 23, Created)
(TrEMBLrel. 23, Last sequence update)
(TrEMBLrel. 24, Last annotation update)
pothetical protein.

s agalactiae (serotype V).
rmicutes; Lactobacillales; Streptococcaceae;
s.
```

```
OX   NCBI_TaxID=216466;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2603 V/R / Serotype V;
RX   MEDLINE=22222988; PubMed=12200547;
RA   Tettelin H., Masignani V., Cieslewicz M.J., Eisen J.A., Peterson
RA   Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
RA   Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C
RA   DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA   Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,
RA   Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mo:
RA   Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Maio:
RA   Rinaudo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
RA   Fraser C.M.;
RT   "Complete genome sequence and comparative genomic analysis of an
RT   emerging human pathogen, serotype V Streptococcus agalactiae.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
DR   EMBL; AE014242; AAM99977.1; -.
DR   TIGR; SAG1096; -.
DR   InterPro; IPR008172; Adenylate_cyc.
DR   Pfam; PF01928; CYTH; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE  190 AA;  22178 MW;  AB2AD33C2CB6FBF3 CRC64;
```

```
Query Match           2.8%;  Score 8;  DB 16;  Length 190;
Best Local Similarity  100.0%;  Pred. No. 27;
Matches    8;  Conservative  0;  Mismatches  0;  Indels  0;

Qy    258 SLRIRTLP 265
          ||||||||
Db     50 SLRIRTLP 57
```

```
RESULT 17
Q8L4K2
ID   Q8L4K2;         PRELIMINARY;         PRT;   193 AA.
AC   Q8L4K2;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
GN   OSJNBA0079H13.8 OR OSJNBB0038H12.21.
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyt
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Kim M.M.,
RA   Overton II L.L., Bera J.J., Tsitrin T., Krol M.I., Jarrahi B.B.,
RA   Jin S.S., Koo H., Zismann V., Hsiao J., Blunt S., Vanaken S.S.,
RA   Utterback T.T., Feldblyum T.V., Yang Q.Q., Haas B.J., Suh B.B.,
RA   Peterson J.J., Quackenbush J., White O., Salzberg S.L., Fraser C
RT   "Oryza sativa chromosome 10 BAC OSJNBA0079H13 genomic sequence."
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M
RA   Overton II.L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,
RA   Fadrosh D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S
RA   Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
RA   Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA   White O., Salzberg S.L., Fraser C.M.;
RT   "Oryza sativa chromosome 10 BAC OSJNBb0038H12 genomic sequence."
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   The Rice Chromosome 10 Sequencing Consortium;
RT   "In-depth view of structure, activity, and evolution of rice
```

```
";
66-1569(2003).

N.A.
ng R.A., McCombie W.R., Messing J., Yuan Q.;
-2003) to the EMBL/GenBank/DDBJ databases.
; AAM54153.1; -.
; AAN04965.1; -.
; AAP52530.1; -.
2; -.
rotein.
. AA; 20812 MW; 719E44FBFBC9A0790 CRC64;
    2.8%;  Score 8;  DB 10;  Length 193;
rity 100.0%;  Pred. No. 27;
nservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

RGE 50
|||
RGE 123


RELIMINARY;      PRT;    197 AA.

rEMBLrel. 12, Created)
rEMBLrel. 12, Last sequence update)
rEMBLrel. 24, Last annotation update)
rotein TM0469.

itima.
motoga; Thermotogales; Thermotogaceae; Thermotoga.
6;

N.A.
DSM 3109;
16; PubMed=10360571;
layton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
key E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
tterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
Smith H.O., Venter J.C., Fraser C.M.;
lateral gene transfer between Archaea and Bacteria from
e of Thermotoga maritima.";
-329(1999).
; AAD35553.1; -.
72374.

rotein; Complete proteome.
. AA; 22919 MW; 41E2C8F3C09180BC CRC64;
    2.8%;  Score 8;  DB 16;  Length 197;
rity 100.0%;  Pred. No. 28;
nservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

FEIS 9
|||
FEIS 142


RELIMINARY;      PRT;    211 AA.

rEMBLrel. 13, Created)
rEMBLrel. 22, Last sequence update)
rEMBLrel. 24, Last annotation update)
in.

lanogaster (Fruit fly).
```

```
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dun
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischman
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA   Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y
RA   Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA   Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA   Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.
RA   Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA   Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA   Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA   McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunco J
RA   Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA   Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.
RA   Stapleton M., Strong R., Svirskas R., Tector C. Tyler D.,
RA   Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT   "Sequencing of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell
RA   Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA   Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA   Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.
RA   Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA   Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA   Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT   "Annotation of Drosophila melanogaster genome."
```

DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Membrane protein, putative.
GN   BRA0991.
OS   Brucella suis.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Brucellaceae; Brucella.
OX   NCBI_TaxID=29461;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1330 / Biovar 1;
RX   MEDLINE=22247741; PubMed=12271122;
RA   Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.
RA   Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,
RA   Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
RA   Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken
RA   Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
RA   Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
RT   "The Brucella suis genome reveals fundamental similarities betwe
RT   animal and plant pathogens and symbionts.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
DR   EMBL; AE014592; AAN34160.1; -.
DR   TIGR; BRA0991; -.
DR   InterPro; IPR007916; UPF0191.
DR   Pfam; PF05252; UPF0191; 1.
KW   Complete proteome.
SQ   SEQUENCE 220 AA; 24796 MW; AC2C06043169497 CRC64;

Query Match   2.8%; Score 8; DB 16; Length 220;
Best Local Similarity   100.0%; Pred. No.30;
Matches   8; Conservative   0; Mismatches   0; Indels   0;

Qy   54 ALLVPLAL 61
          ||||||||
Db   132 ALLVPLAL 139

RESULT 22
Q89GW9
ID   Q89GW9
AC   Q89GW9;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Blr6226 protein.
GN   BLR6226.
OS   Bradyrhizobium japonicum.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Bradyrhizobiaceae; Bradyrhizobium.
OX   NCBI_TaxID=375;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=USDA 110;
RX   MEDLINE=22484998; PubMed=12597275;
RA   Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA   Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA   Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada
RA   Tabata S.;
RT   "Complete genomic sequence of nitrogen-fixing symbiotic bacteriu
RT   Bradyrhizobium japonicum USDA110.";
RL   DNA Res. 9:189-197(2002).
DR   EMBL; AP005957; BAC51491.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR000515; BPD_transp.
DR   Pfam; PF00528; BPD_transp; 1.
KW   Complete proteome.
SQ   SEQUENCE 232 AA; 23704 MW; CD805BD1F43F1B46 CRC64;

Query Match   2.8%; Score 8; DB 16; Length 232;
Best Local Similarity   100.0%; Pred. No.32;
Matches   8; Conservative   0; Mismatches   0; Indels   0;

R-2000) to the EMBL/GenBank/DDBJ databases.
N.A.
elhiker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
R-2000) to the EMBL/GenBank/DDBJ databases.
N.A.

P-2002) to the EMBL/GenBank/DDBJ databases.
7; AAF49452.2; -.
0036638; CG13033.
004011; Gyr.
004019; YLP_motif.
; GYR; 1.
; YLP; 5.
1 AA; 23779 MW; D2554983E91F5107 CRC64;

arity   100.0%; Score 8; DB 5; Length 211;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

VVSL 79
||||
VVSL 16

PRELIMINARY;   PRT;   211 AA.

TrEMBLrel. 21, Created)
TrEMBLrel. 21, Last sequence update)
TrEMBLrel. 24, Last annotation update)

lanogaster (Fruit fly).
tazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
opterygota; Diptera; Brachycera; Muscomorpha;
Drosophilidae; Drosophila.
27;

N.A.
ey;
Brokstein P., Hong L., Agbayani A., Carlson J.,
avez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
nzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
fungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
uanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,

R-2002) to the EMBL/GenBank/DDBJ databases.
4; AAL90372.1; -.
0063673; BcDNA:RE50345.
004011; Gyr.
004019; YLP_motif.
; GYR; 1.
; YLP; 5.
1 AA; 23780 MW; 82FF4983E91F510A CRC64;

arity   100.0%; Score 8; DB 5; Length 211;
Conservative   0; Mismatches   0; Indels   0; Gaps   0;

VVSL 79
||||
VVSL 16

PRELIMINARY;   PRT;   220 AA.

(TrEMBLrel. 23, Created)
(TrEMBLrel. 23, Last sequence update)

LGL 65
|||
LGL 209

ELIMINARY; PRT; 276 AA.

rEMBLrel. 24, Created)
rEMBLrel. 24, Last sequence update)
rEMBLrel. 25, Last annotation update)

aliana (Mouse-ear cress).
idiplantae; Streptophyta; Embryophyta; Tracheophyta;
Magnoliophyta; eudicotyledons; core eudicots; rosids;
rassicales; Brassicaceae; Arabidopsis.
2;

N.A.
.H., Kim C.J., Shinn P., Bowser L., Carninci P.,
ashizaki Y., Hsuan V.W., Ishida J., Jones T., Kamiya A.,
.G., Kawai J., Lam B., Lin J., Miranda M., Narusaka M.,
dera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
wick A., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G.,
zaki K., Davis R.W., Theologis A., Ecker J.R.;
RF clones.";
. AAP21226.1; -.
; F:electron transporter activity; IEA.
; P:electron transport; IEA.
00345; CytC_heme_BS.
01841; Znf_ring.
zf-C3HC4_1.
; RING; 1.
90; CYTOCHROME C; 1.
89; ZF_RING_2; 1.
i AA; 31368 MW; 99DE3DA0CB2C0BDF CRC64;

rity 2.8%; Score 8; DB 10; Length 276;
100.0%; Pred. No. 37;
onservative 0; Mismatches 0; Indels 0; Gaps 0;

HISA 10
|||
HISA 221

ELIMINARY; PRT; 278 AA.

rEMBLrel. 10, Created)
rEMBLrel. 10, Last sequence update)
rEMBLrel. 24, Last annotation update)
protein AMPE.

eruginosa.
eobacteria; Gammaproteobacteria; Pseudomonadales;
ae; Pseudomonas.
;

N.A.
692 / PAO1;
Dargis M., Huletsky A.;
in Pseudomonas aeruginosa encodes a negative regulator
lactamase expression.";
-1998) to the EMBL/GenBank/DDBJ databases.

eruginosa.
eobacteria; Gammaproteobacteria; Pseudomonadales;
ae; Pseudomonas.

N.A.
692 / PAO1;
337; PubMed=10984043;
ham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,

RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AF082575; AAC98784.1; -.
DR EMBL; AE004866; AAG07909.1; -.
DR PIR; D83080; D83080.
DR GO; GO:0016021; C:integral to membrane; IEA.
KW Transmembrane; Complete proteome.
SQ SEQUENCE 278 AA; 30793 MW; C623F1AB0691CFEF CRC64;

Query Match 2.8%; Score 8; DB 16; Length 278;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 8; Conservative 0; Mismatches 0; Indels 0; G

Qy 70 LGLLLAVV 77
||||||||
Db 47 LGLLLAVV 54

RESULT 25
Q82J43
ID Q82J43 PRELIMINARY; PRT; 278 AA.
AC Q82J43;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative metalloproteinase.
GN SAV2939.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industr
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AP005033; BAC70650.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008237; F:metallopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000013; Peptidase_M7.
DR Pfam; PF02031; Peptidase_M7; 1.
DR PRINTS; PR00787; NEUTRALPTASE.
DR ProDom; PD016028; Peptidase_M7; 1.
KW Complete proteome.
SQ SEQUENCE 278 AA; 28113 MW; 9545813BCAC0BFA2 CRC64;

Query Match 2.8%; Score 8; DB 16; Length 278;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 8; Conservative 0; Mismatches 0; Indels 0; G

Qy 59 LALGLGLA 66

||||
HLGLA 20

PRELIMINARY;   PRT;   279 AA.

(TrEMBLrel. 17, Created)
(TrEMBLrel. 17, Last sequence update)
(TrEMBLrel. 23, Last annotation update)
k protein (Mage-g1) (RIKEN cDNA 5730494G16 gene).
K.
(Mouse).
etazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
theria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
0090;

M.N.A.
/6J; TISSUE=Medulla oblongata;
5660; PubMed=11217851;
inagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
zawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
azaki T., Gojobori T., Bono H., Kasukawa T., Saito R.,
atsuda H.A., Ashburner M., Batalov S., Casavant T.,
v., Gaasterland T., Gissi C., King B., Kochiwa H.,
vis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
ido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
ffelli D., Bojunga N., Carninci P., de Bonaldo M.,
J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
chionni L., Mashima J., Mazzarelli J., Mombaerts P.,
Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
ato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
oyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
.,;
annotation of a full-length mouse cDNA collection.";
85-690(2001).
50; BAB31133.1; -.
3897; 5730494G16Rik.
R002190; MAGE.
4; MAGE; 1.
0838; MAGE; 1.
79 AA; 31474 MW; 5E243590A99F15F0 CRC64;

larity   2.8%;   Score 8;   DB 11;   Length 279;
Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

ASLSA 87
       ||||||
ASLSA 48

PRELIMINARY;   PRT;   279 AA.

(TrEMBLrel. 17, Created)
(TrEMBLrel. 17, Last sequence update)
(TrEMBLrel. 23, Last annotation update)
k protein (Mage-g1) (RIKEN cDNA 5730494G16 gene).
K.
(Mouse).
etazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
theria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
0090;

M.N.A.
/6J; TISSUE=Embryo, and Embryonic stem cells;

RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbus
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washi
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilmi
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Auquier P.H., Chomez P.M., De Backer O.R., Bertrand M.J.M.;
RT   "Ten new murine members of the MAGE gene family.";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Mammary gland;
RA   Strausberg R.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Head, and Spinal cord;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium;
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotat
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK017727; BAB30899.1; -.
DR   EMBL; AK010294; BAB26830.1; -.
DR   EMBL; AF319979; AAK01207.1; -.
DR   EMBL; BC034892; AAH34892.1; -.
DR   EMBL; AK049759; BAC33907.1; -.
DR   EMBL; AK076471; BAC36358.1; -.
DR   MGD; MGI:1913897; 5730494G16Rik.
DR   InterPro; IPR002190; MAGE.
DR   Pfam; PF01454; MAGE; 1.
DR   PROSITE; PS50838; MAGE; 1.
SQ   SEQUENCE   279 AA;   31460 MW;   FE243519BD63160 CRC64;

Query Match   2.8%;   Score 8;   DB 11;   Length 279;
Best Local Similarity   100.0%;   Pred. No. 38;
Matches   8;   Conservative   0;   Mismatches   0;   Indels   0;

Qy   80 GSRASLSA 87
              ||||||||
Db   41 GSRASLSA 48

RESULT 28
Q9VNP0
ID   Q9VNP0   PRELIMINARY;   PRT;   306 AA.
AC   Q9VNP0; Q8SZB8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   CG1169 protein (RE07882p).
GN   CG1169.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

...osophilidae; Drosophila.
7;

N.A.
...y;
06; PubMed=10731112;
...lniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
..., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
...ewis S.E., Richards S., Ashburner M., Henderson S.N.,
...ortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
...e C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
...bayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
...enos P.V., Berman B.P., Bhandari D., Bolshakov S.,
...tchan M.R., Bouck J., Brokstein P., Brottier P.,
...usam D.A., Butler H., Cadieu E., Center A., Chandra I.,
...awley S., Dahlke C., Davenport L.B., Davies P.,
Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
...p L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
...vangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
...rielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
...g F., Gorrell J.H., Gu Z., Guan P., Harris M.,
...arvey D., Heiman T.J., Hernandez J.R., Houck J.,
...ston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
...ush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
...odira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
...Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
...B., McIntosh T.C., McLeod M.P., McPherson D.,
...ilshina N.V., Mobarry C., Morris J., Moshrefi A.,
...y M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
...elson K.A., Nixon K., Nusskern D.R., Pacleb J.M., Nelson D.L.,
...Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
...mington K., Saunders R.D.C., Scheeler F., Shen H.,
...en-Kiamos I., Simpson M., Skupski M.P., Smith T.,
...dling A.C., Stapleton M., Strong R., Sun E.,
...ector C., Turner R., Venter E., Wang A.H., Wang X.,
...assarman D.A., Weinstock G.M., Weissenbach J.,
Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
...F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
...ong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
...ers E.W., Rubin G.M., Venter J.C.";
...quence of Drosophila melanogaster.";
...85-2195(2000).

N.A.
Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
...cayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
...Center A., Champe M., Davenport L.B., Dietz S.M.,
...sett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
...rise E., Galle R.F., Garg N.S., George R.A.,
...louck J., Hoskins R.A., Hostin D., Howland T.J.,
...lali M., Kruse D., Li P., Mattei B., Moshrefi A.,
...Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
...ragas V., Park S., Patel S., Pfeiffer B., Scheeler F.,
...S., Pittman G.S., Puri V., Richards S., Scheeler F.,
...Strong R., Svirskas R., Tector C., Tyler D.,
...Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
...: Drosophila melanogaster genome.";
...R-2000) to the EMBL/GenBank/DDBJ databases.

N.A.
...aby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
...huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
...cgman C., Berman B., Carlson J.W., Celniker S.E.,
...sdale R., Emmert D., Frise E., de Grey A., Harris N.,
..., Marshall B., Millburn G., Richter J., Russo S.,
..., Smith E., Shu S., Smutniak F., Whitfield E.,
...Gelbart W.M., Shu S., Rubin G.M., Mungall C.J., Lewis S.E.;
...f Drosophila melanogaster genome.";
...R-2000) to the EMBL/GenBank/DDBJ databases.

RP   SEQUENCE FROM N.A.
RA   Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA   George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA   Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.
RA   Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.
RA   Celniker S.;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003600; AAF51889.2; -.
DR   EMBL; AY070982; AAL48604.1; -.
DR   FlyBase; FBgn0037428; CG1169.
SQ   SEQUENCE   306 AA; 34083 MW; 32B69371475A48F9 CRC64;

Query Match   2.8%; Score 8; DB 5; Length 306;
Best Local Similarity   100.0%; Pred. No. 41;
Matches   8; Conservative   0; Mismatches   0; Indels   0; G

Qy   56 LVPLALGL 63
              ||||||||
Db   170 LVPLALGL 177

RESULT 29
Q864L1
ID   Q864L1   PRELIMINARY;   PRT;   317 AA.
AC   Q864L1;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Melanocortin-1 receptor.
GN   MC1R.
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=3;
RX   MEDLINE=22572539; PubMed=12687585;
RA   Mundy N.I., Kelly J.;
RT   "Evolution of a pigmentation gene, the melanocortin-1 receptor, i
RT   primates.";
RL   Am. J. Phys. Anthropol. 121:67-80(2003).
DR   EMBL; AY205086; AAP30960.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   317 AA; 34710 MW; 8815D21464BD2475 CRC64;

Query Match   2.8%; Score 8; DB 6; Length 317;
Best Local Similarity   100.0%; Pred. No. 42;
Matches   8; Conservative   0; Mismatches   0; Indels   0; G

Qy   137 RARRAIAA 144
              ||||||||
Db   160 RARRAIAA 167

DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   317 AA;   34782 MW;   96A807F178FEC21 CRC64;

Query Match
Best Local Similarity   2.8%;
Matches   8; Conservative   0; Mismatches   0; Indels   0;

Qy   137 RARRAIAA 144
             ||||||||
Db   160 RARRAIAA 167

RESULT 32
Q864J7
ID   Q864J7   PRELIMINARY;   PRT;   317 AA.
AC   Q864J7;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Melanocortin-1 receptor.
GN   MC1R.
OS   Macaca silenus (Lion-tailed macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=54601;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1;
RX   MEDLINE=22572539; PubMed=12687585;
RA   Mundy N.I., Kelly J.;
RT   "Evolution of a pigmentation gene, the melanocortin-1 receptor,
RT   primates.";
RL   Am. J. Phys. Anthropol. 121:67-80(2003).
DR   EMBL; AY205100; AAP30974.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. .
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   317 AA;   34782 MW;   96A807F178FEC21 CRC64;

Query Match
Best Local Similarity   2.8%;   100.0%;   Pred. No. 42;
Matches   8; Conservative   0; Mismatches   0; Indels   0;

Qy   137 RARRAIAA 144
             ||||||||
Db   160 RARRAIAA 167

RESULT 33
Q864J5
ID   Q864J5   PRELIMINARY;   PRT;   317 AA.
AC   Q864J5;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Melanocortin-1 receptor.
GN   MC1R.
OS   Macaca nigra (Celebes black macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=54600;

RESULT 35
Q864G7
```
ID   Q864G7      PRELIMINARY;    PRT;   317 AA.
AC   Q864G7;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Melanocortin-1 receptor.
GN   MC1R.
OS   Ateles paniscus (Black spider monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Atelinae; Ate.
OX   NCBI_TaxID=9510;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=3;
RX   MEDLINE=22572539; PubMed=12687585;
RA   Mundy N.I., Kelly J.;
RT   "Evolution of a pigmentation gene, the melanocortin-1 receptor, in
RT   primates.";
RL   Am. J. Phys. Anthropol. 121:67-80(2003).
DR   EMBL; AY205130; AAP31004.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin.  .
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   317 AA;   34719 MW;   5481D6A1B9085D43 CRC64;

Query Match             2.8%;
Best Local Similarity   100.0%;  Score 8;  DB 6;  Length 317;
Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    137 RARRAIAA 144
          ||||||||
Db    160 RARRAIAA 167
```

RESULT 36
Q864G6
```
ID   Q864G6      PRELIMINARY;    PRT;   317 AA.
AC   Q864G6;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Melanocortin-1 receptor.
GN   MC1R.
OS   Alouatta seniculus (Red howler monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Alouattinae;
OC   Alouatta.
OX   NCBI_TaxID=9503;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=471;
RX   MEDLINE=22572539; PubMed=12687585;
RA   Mundy N.I., Kelly J.;
RT   "Evolution of a pigmentation gene, the melanocortin-1 receptor, in
RT   primates.";
RL   Am. J. Phys. Anthropol. 121:67-80(2003).
DR   EMBL; AY205131; AAP31005.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
```

```
RP   SEQUENCE FROM N.A.
RC   STRAIN=2;
RX   MEDLINE=22572539; PubMed=12687585;
RA   Mundy N.I., Kelly J.;
RT   "Evolution of a pigmentation gene, the melanocortin-1 receptor, :
RT   primates.";
RL   Am. J. Phys. Anthropol. 121:67-80(2003).
DR   EMBL; AY205134; AAP31008.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. .
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   317 AA;  34692 MW;  91320B374CDB75DB CRC64;

Query Match
Best Local Similarity   2.8%;  Score 8; DB 6;  Length 317;
Matches    8; Conservative   0; Mismatches   0;  Indels   0;

Qy   137 RARRAIAA 144
         ||||||
Db   160 RARRAIAA 167


RESULT 39
Q864G2
ID   Q864G2      PRELIMINARY;       PRT;   317 AA.
AC   Q864G2;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Melanocortin-1 receptor.
GN   MC1R.
OS   Alouatta palliata (Mantled howler monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
OC   Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Alouattinae;
OC   Alouatta.
OX   NCBI_TaxID=30589;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1;
RX   MEDLINE=22572539; PubMed=12687585;
RA   Mundy N.I., Kelly J.;
RT   "Evolution of a pigmentation gene, the melanocortin-1 receptor,
RT   primates.";
RL   Am. J. Phys. Anthropol. 121:67-80(2003).
DR   EMBL; AY205135; AAP31009.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. .
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   317 AA;  34670 MW;  D5414E350E2435DE CRC64;

Query Match
Best Local Similarity   2.8%;  Score 8; DB 6;  Length 317;
Matches    8; Conservative   0; Mismatches   0;  Indels   0;

Qy   137 RARRAIAA 144
         ||||||
Db   160 RARRAIAA 167
```

```
237; G_PROTEIN_RECEP_F1_1; 1.
262; G_PROTEIN_RECEP_F1_2; 1.
7 AA;  34830 MW;  87F7EFAE347671E4 CRC64;

         2.8%;  Score 8; DB 6;  Length 317;
arity 100.0%;  Pred. No. 42;
onservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

RAIAA 144
||||
RAIAA 167


PRELIMINARY;       PRT;   317 AA.

(TrEMBLrel. 24, Created)
(TrEMBLrel. 24, Last sequence update)
(TrEMBLrel. 25, Last annotation update)
-1 receptor.

(Bolivian red howler monkey).
etazoa; Chordata; Craniata; Vertebrata; Euteleostomi
theria; Primates; Platyrrhini; Cebidae; Alouattinae;

21123;

4 N.A.

2539; PubMed=12687585;
Kelly J.;
= a pigmentation gene, the melanocortin-1 receptor, in
Anthropol. 121:67-80(2003).
33; AAP31007.1; -.
21; C:integral to membrane; IEA.
72; F:receptor activity; IEA.
34; F:rhodopsin-like receptor activity; IEA.
86; P:G-protein coupled receptor protein signalin. . .; IEA.
R000276; GPCR_Rhodpsn.
1; 7tm_1; 1.
237; GPCRRHODOPSN.
237; G_PROTEIN_RECEP_F1_1; 1.
262; G_PROTEIN_RECEP_F1_2; 1.
17 AA;  34686 MW;  BA7E14AFEC7EA971 CRC64;

         2.8%;  Score 8; DB 6;  Length 317;
larity 100.0%;  Pred. No. 42;
Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

RAIAA 144
||||
RAIAA 167


PRELIMINARY;       PRT;   317 AA.

(TrEMBLrel. 24, Created)
(TrEMBLrel. 24, Last sequence update)
(TrEMBLrel. 25, Last annotation update)
-1 receptor.

aya (Black howler monkey).
etazoa; Chordata; Craniata; Vertebrata; Euteleostomi
theria; Primates; Platyrrhini; Cebidae; Alouattinae;

502;
```

DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   317 AA; 34714 MW; C1F5DA35032717D7 CRC64;

Query Match   2.8%; Score 8; DB 6; Length 317;
Best Local Similarity   100.0%; Pred. No. 42;
Matches   8; Conservative   0; Mismatches   0; Indels   0; G

Qy   137 RARRAIAA 144
          ||||||||
Db   160 RARRAIAA 167

RESULT 42
Q864F7
ID   Q864F7   PRELIMINARY;   PRT;   317 AA.
AC   Q864F7;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Melanocortin-1 receptor.
GN   MC1R.
OS   Varecia variegata variegata.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Strepsirhini; Lemuridae; Varecia.
OX   NCBI_TaxID=87289;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1;
RX   MEDLINE=22572539; PubMed=12687585;
RA   Mundy N.I., Kelly J.;
RT   "Evolution of a pigmentation gene, the melanocortin-1 receptor, in
RT   primates.";
RL   Am. J. Phys. Anthropol. 121:67-80(2003).
DR   EMBL; AY205140; AAP31014.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. .
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   317 AA; 34714 MW; C1F5DA35032717D7 CRC64;

Query Match   2.8%; Score 8; DB 6; Length 317;
Best Local Similarity   100.0%; Pred. No. 42;
Matches   8; Conservative   0; Mismatches   0; Indels   0; G

Qy   137 RARRAIAA 144
          ||||||||
Db   160 RARRAIAA 167

RESULT 43
Q864F6
ID   Q864F6   PRELIMINARY;   PRT;   317 AA.
AC   Q864F6;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Melanocortin-1 receptor.
GN   MC1R.
OS   Eulemur fulvus (brown lemur).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Strepsirhini; Lemuridae; Eulemur.
OX   NCBI_TaxID=13515;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2;

AC   Q864F4;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Melanocortin-1 receptor.
GN   MC1R.
OS   Lemur catta (Ring-tailed lemur).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
OC   Mammalia; Eutheria; Primates; Strepsirhini; Lemuridae; Lemur.
OX   NCBI_TaxID=9447;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=3;
RX   MEDLINE=22572539; PubMed=12687585;
RA   Mundy N.I., Kelly J.;
RT   "Evolution of a pigmentation gene, the melanocortin-1 receptor,
RT   primates.";
RL   Am. J. Phys. Anthropol. 121:67-80(2003).
DR   EMBL; AY205143; AAP31017.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. .;
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   317 AA; 34667 MW; 3E7419FDEC2DE738 CRC64;

Query Match   2.8%; Score 8; DB 6; Length 317;
Best Local Similarity   100.0%; Pred. No. 42;
Matches   8; Conservative   0; Mismatches   0; Indels   0;

Qy   137 RARRAIAA 144
              ||||||||
Db   160 RARRAIAA 167

RESULT 46
O58554
ID   O58554   PRELIMINARY;   PRT;   339 AA.
AC   O58554;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-JAN-1999 (TrEMBLrel. 09, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein PH0824.
GN   PH0824.
OS   Pyrococcus horikoshii.
OC   Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococca
OC   Pyrococcus.
OX   NCBI_TaxID=53953;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OT3;
RX   MEDLINE=98344137; PubMed=9679194;
RA   Kawarabayasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA   Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Naga
RA   Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku
RA   Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguc
RA   Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA   Masuchi Y., Shizuya H., Kikuchi H.;
RT   "Complete sequence and gene organization of the genome of a hype
RT   thermophilic archaebacterium, Pyrococcus horikoshii OT3.";
RL   DNA Res. 5:55-76(1998).
DR   EMBL; AP000003; BAA29917.1; -.
DR   PIR; C71132; C71132.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008508; F:bile acid:sodium symporter activity; IEA.
DR   GO; GO:0006814; P:sodium ion transport; IEA.
DR   InterPro; IPR002657; BilAc/Na_symport.
DR   Pfam; PF01758; SBF; 1.

539; PubMed=12687585;
elly J.;
a pigmentation gene, the melanocortin-1 receptor, in

Anthropol. 121:67-80(2003).
1; AAP31015.1; -.
21; C:integral to membrane; IEA.
72; F:receptor activity; IEA.
84; F:rhodopsin-like receptor activity; IEA.
86; P:G-protein coupled receptor protein signalin. .; IEA.
000276; GPCR_Rhodpsn.
.; 7tm_1; 1.
237; GPCRRHODOPSN.
0237; G_PROTEIN_RECEP_F1_1; 1.
0262; G_PROTEIN_RECEP_F1_2; 1.

.7 AA; 34748 MW; DAF913C3B9ECC2AF CRC64;

larity   2.8%; Score 8; DB 6; Length 317;
           100.0%; Pred. No. 42;
Conservative   0; Mismatches   0; Indels   0; Gaps   0;

RAIAA 144
|||||
RAIAA 167

PRELIMINARY;   PRT;   317 AA.

(TrEMBLrel. 24, Created)
(TrEMBLrel. 24, Last sequence update)
(TrEMBLrel. 25, Last annotation update)
-1 receptor.

useus (Gray gentle lemur) (Bamboo lemur).
etazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
heria; Primates; Strepsirhini; Lemuridae; Hapalemur.
3557;

4 N.A.

2539; PubMed=12687585;
Kelly J.;
= a pigmentation gene, the melanocortin-1 receptor, in
-1 receptor.

Anthropol. 121:67-80(2003).
12; AAP31016.1; -.
21; C:integral to membrane; IEA.
72; F:receptor activity; IEA.
84; F:rhodopsin-like receptor activity; IEA.
86; P:G-protein coupled receptor protein signalin. .; IEA.
R000276; GPCR_Rhodpsn.
.; 7tm_1; 1.
237; GPCRRHODOPSN.
0237; G_PROTEIN_RECEP_F1_1; 1.
0262; G_PROTEIN_RECEP_F1_2; 1.

17 AA; 34424 MW; 7FB912CCC3F5EC71 CRC64;

larity   2.8%; Score 8; DB 6; Length 317;
           100.0%; Pred. No. 42;
Conservative   0; Mismatches   0; Indels   0; Gaps   0;

RAIAA 144
|||||
RAIAA 167

PRELIMINARY;   PRT;   317 AA.

rotein; Complete proteome.
AA; 37228 MW; E91697D5C8C3705F CRC64;

rity　　2.8%; Score 8; DB 17; Length 339;
nservative　0; Mismatches　0; Indels　0; Gaps　0;

Qy　　60 ALGLGLAL 67
　　　　|||||||||
Db　　311 ALGLGLAL 318

```
AVV  77
     |||
AVV 118
```

RESULT 49

ELIMINARY;　　PRT;　　342 AA.

Q8TUU8
ID   Q8TUU8　　　　　PRELIMINARY;　　　　PRT;　　370 AA.
AC   Q8TUU8;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Permease subunit of a ABC-type transport system involved in
DE   lipoprotein release.
GN   MK1655.
OS   Methanopyrus kandleri.
OC   Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyrace
OC   Methanopyrus.
OX   NCBI_TaxID=2320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AV19 / DSM 6324 / JCM 9639;
RX   MEDLINE=21927647; PubMed=11930014;
RA   Slesarev A.I., Mezheva K.V., Makarova K.S., Polushin N.N.,
RA   Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA   Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA   Malykh A.G., Koonin E.V., Kozyavkin S.A.,
RT   "The complete genome of hyperthermophile Methanopyrus kandleri AV
RT   and monophyly of archaeal methanogens.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
DR   EMBL; AE010455; AAM02868.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR003838; DUF214.
DR   Pfam; PF02687; FtsX; 1.
KW   Complete proteome.
SQ   SEQUENCE 370 AA; 39411 MW; B07662EA1E5A644E CRC64;

ELIMINARY;　　PRT;　　343 AA.

rEMBLrel. 03, Created)
rEMBLrel. 03, Last sequence update)
rEMBLrel. 24, Last annotation update)

Human).
azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
eria; Primates; Catarrhini; Hominidae; Homo.
6;

N.A.
28; PubMed=8702217;
on N., Lee S.W., Liebert M., Grossman H.B.;
lysis of a gene, BB1, overexpressed in bladder and
ma.";
. 16:1085-1090(1996).
AAB37433.1; -.
04299; MBOAT_fam.
MBOAT; 1.
AA; 38163 MW; 2B479EA8CFF1B91C CRC64;

Query Match　　　　　　2.8%; Score 8; DB 17; Length 370;
Best Local Similarity　100.0%; Pred. No. 49;
Matches　　8; Conservative　0; Mismatches　0; Indels　0; G

rity　　2.8%; Score 8; DB 4; Length 342;
nservative　0; Mismatches　0; Indels　0; Gaps　0;

Qy　　61 LGLGLALA 68
　　　　||||||||
Db　　336 LGLGLALA 343

```
LAL  67
     |||
LAL 317
```

RESULT 50

ELIMINARY;　　PRT;　　343 AA.

O53860
ID   O53860　　　　　PRELIMINARY;　　　　PRT;　　372 AA.
AC   O53860;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein CysM3
GN   CYSM3 OR RV0848 OR MTV043.41.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37Rv;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harri
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the

rEMBLrel. 17, Created)
rEMBLrel. 17, Last sequence update)
rEMBLrel. 22, Last annotation update)
rotein.

Human).
azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
eria; Primates; Catarrhini; Hominidae; Homo.
6;

N.A.
and Colon;

3-2001) to the EMBL/GenBank/DDBJ databases.
; AAH03164.1; -.
; AAH02552.1; -.
04299; MBOAT_fam.
MBOAT; 1.
protein.
AA; 38727 MW; F71E7DBF74BD9BB7 CRC64;

rity　　100.0%; Pred. No. 45; DB 4; Length 343;
nservative　0; Mismatches　0; Indels　0; Gaps　0;

RP   SEQUENCE FROM N.A.
RC   STRAIN=LL/S2;
RX   MEDLINE=22557897; PubMed=12670979;
RA   Wood G.E., Haydock A.K., Leigh J.A.;
RT   "Function and Regulation of the Formate Dehydrogenase Genes of the
RT   Methanogenic Archaeon Methanococcus maripaludis.";
RL   J. Bacteriol. 185:2548-2554(2003).
DR   EMBL; AY236516; AAO85931.1; -.
DR   GO; GO:0005489; F:electron transporter activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001450; 4Fe4S_ferredoxin.
DR   InterPro; IPR007525; FrhB_FdhB_C.
DR   InterPro; IPR007516; FrhB_FdhB_N.
DR   Pfam; PF00037; fer4; 1.
DR   Pfam; PF04432; FrhB_FdhB_C; 1.
DR   Pfam; PF04422; FrhB_FdhB_N; 1.
DR   PROSITE; PS00198; 4Fe4S_FERREDOXIN; 2.
SQ   SEQUENCE 375 AA; 42431 MW; AB096789.18AEC8AD CRC64;

Query Match   2.8%; Score 8; DB 1; Length 375;
Best Local Similarity   100.0%; Pred. No. 49;
Matches   8; Conservative   0; Mismatches   0; Indels   0; C

Qy   217 LVDGVLAL 224
               ||||||||
Db   35 LVDGVLAL 42

RESULT 53
Q9KYT3
ID   Q9KYT3   PRELIMINARY;   PRT;   377 AA.
AC   Q9KYT3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Putative integral membrane protein.
GN   SCO5682 OR SC5H4.06C.
OS   Streptomyces coelicolor.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2) / M145;
RX   MEDLINE=21996410; PubMed=12000953;
RA   Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA   Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA   Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA   Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA   Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA   Rabbinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA   Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor
RA   Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA   Hopwood D.A.;
RT   "Complete genome sequence of the model actinomycete Streptomyces
RT   coelicolor A3(2).";
RL   Nature 417:141-147(2002).
DR   EMBL; AL939124; CAB91118.1; -.
KW   Complete proteome.
SQ   SEQUENCE 377 AA; 37614 MW; A35DA0437F044C46 CRC64;

Query Match   2.8%; Score 8; DB 16; Length 377;
Best Local Similarity   100.0%; Pred. No. 50;
Matches   8; Conservative   0; Mismatches   0; Indels   0; C

Qy   59 LALGLGLA 66
               ||||||||
Db   327 LALGLGLA 334

RESULT 54
Q9SHD8
ID   Q9SHD8   PRELIMINARY;   PRT;   387 AA.

---

me sequence.";
7-544(1998).
4; CAA17654.1; -.
H70813.
1JBQ.
Rv0848; -.
9; F:lyase activity; IEA.
0; P:amino acid metabolism; IEA.
001926; B6_enzyme_beta.
; PALP; 1.
protein; Complete proteome.
2 AA; 40118 MW; 927386BE1DF5FB6C CRC64;

arity   2.8%; Score 8; DB 16; Length 372;
100.0%; Pred. No. 49;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

LALA 68
||||
LALA 103

PRELIMINARY;   PRT;   372 AA.

(TrEMBLrel. 25, Created)
(TrEMBLrel. 25, Last sequence update)
(TrEMBLrel. 25, Last annotation update)
eine synthase A CYSK2 (O-acetylserine sulfhydrylase)
ne (Thiol)-lyase) (CSASE) (EC 2.5.1.47).
71.
bovis.
inobacteria; Actinobacteridae; Actinomycetales;
neae; Mycobacteriaceae; Mycobacterium.
65;

N.A.
/97;
107; PubMed=12788972;
iglmeier K., Camus J.-C., Medina N., Mansoor H.,
hoy S., Grondin S., Lacroix C., Monsempe C., Simon S.,
kin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
genome sequence of Mycobacterium bovis.";
cad. Sci. U.S.A. 100:7877-7882(2003).
6; CAD93733.1; -.
Complete proteome.
2 AA; 40106 MW; 6FD459DFA6FB4284 CRC64;

arity   2.8%; Score 8; DB 16; Length 372;
100.0%; Pred. No. 49;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

LALA 68
||||
LALA 103

PRELIMINARY;   PRT;   375 AA.

(TrEMBLrel. 24, Created)
(TrEMBLrel. 24, Last sequence update)
(TrEMBLrel. 25, Last annotation update)
drogenase beta subunit.

maripaludis.
archaeota; Methanococci; Methanococcales;
eae; Methanococcus.
152;

```
DT   rEMBLrel. 13, (Created)
DT   rEMBLrel. 13, Last sequence update)
DT   rEMBLrel. 25, Last annotation update)
DE   1.20 (At2g45000 protein).
GN   T2G45000.
OS   aliana (Mouse-ear cress).
OC   idiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   rassicales; Brassicaceae; Arabidopsis.
OX   2;
RN   N.A.
RA   H., Cheuk R., Kim C.J., Lim J., Meyers M.C., Banh J.,
RA   ninci P., Chang E., Dale J.M., Goldsmith A.D.,
RA   , Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA   B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA   dera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA   wick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA   u G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RT   DNA clones.";
RL   -2002) to the EMBL/GenBank/DDBJ databases.
RN   N.A.
RA   S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA   , Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA   hen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA   lin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA   rusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA   P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RT   ull Length cDNA Clones.";
RL   -2002) to the EMBL/GenBank/DDBJ databases.
RN   N.A.
RA   umbia;
RX   87; PubMed=10617197;
RA   , Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA   son T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA   tchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA   hen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA   rrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA   ., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA   Fraser C.M., Venter J.C.;
RT   analysis of chromosome 2 of the plant Arabidopsis
RL   -768 (1999).
RN   N.A.
DR   umbia;
DR   ; AAL69462.1; -.
DR   ; AAL86303.1; -.
DR   ; AAD32835.1; -.
KW   84885.
SQ   rotein.
SQ   AA; 40584 MW; AF6C6B3BAC9BF69A CRC64;

Query Match       2.8%;  Score 8;  DB 10;  Length 387;
Best Local Similarity  100.0%;  Pred. No. 51;
Matches   8;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    EED 100
       |||
Db    EED 377

RESULT 56
ELIMINARY;    PRT;    397 AA.
```

```
DT   01-JUN-2002 (TrEMBLrel. 21, (Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein Atu3948.
GN   ATU3948 OR AGR_L_1808.
OS   Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX   NCBI_TaxID=176299;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21608550; PubMed=11743193;
RA   Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA   Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA   Chen Y., Paulsen I.T. Eisen J.A., Karp P.D., Bovee D. Sr.,
RA   Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA   Kutyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA   Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon
RA   Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.
RA   Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.
RA   Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA   Nester E.W.,
RT   "The genome of the natural genetic engineer Agrobacterium tumefaci
RT   C58.";
RL   Science 294:2317-2323 (2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21608551; PubMed=11743194;
RA   Goodner B., Hinkle G., Gattung S., Miller N., Blanchard C.,
RA   Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin
RA   Houmiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA   Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz E
RA   Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
RA   Cielo C., Slater S.;
RT   "Genome sequence of the plant pathogen and biotechnology agent
RT   Agrobacterium tumefaciens C58.";
RL   Science 294:2323-2328(2001).
DR   EMBL; AE009325; AAK44750.1; ALT_INIT.
DR   EMBL; AE008289; AAK89478.1; -.
DR   PIR; AH3041; AH3041.
DR   PIR; D98244; D98244.
DR   InterPro; IPR001220; Lectin_legB.
DR   InterPro; IPR001608; UPF0001.
DR   Pfam; PF01168; Ala_racemase_N; 1.
DR   PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE 397 AA; 41708 MW; 70074E32A46AE86 CRC64;

Query Match       2.8%;  Score 8;  DB 16;  Length 397;
Best Local Similarity  100.0%;  Pred. No. 52;
Matches   8;  Conservative  0;  Mismatches  0;  Indels  0;  Ga

Qy    251 LALRPGSS 258
          ||||||||
Db    341 LALRPGSS 348

RESULT 56
Q7U9H4
ID   Q7U9H4              PRELIMINARY;    PRT;    431 AA.
AC   Q7U9H4;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Possible bicarbonate transporter, ICT family.
GN   SYNW0284.
OS   Synechococcus sp. (strain WH8102).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX   NCBI_TaxID=84588;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22825697; PubMed=12917641;
RA   Palenik B., Brahamsha B., Larimer F.W., Land M., Hauser L., Chain
```

RA   Regala W., Allen E.E., McCarren J., Paulsen I.,
RA   Partensky F., Webb E.A., Waterbury J.;
RT   f a motile marine Synechococcus.";
RL   37-1042(2003).
DR   9; CAE06799.1; -.
eome.
1 AA; 46300 MW; 66295F913903DBAE CRC64;

arity   2.8%; Score 8; DB 16; Length 431;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

GLlL 74
      |||
GLlL 371

PRELIMINARY;   PRT;   435 AA.
TrEMBLrel. 20, Created)
TrEMBLrel. 20, Last sequence update)
TrEMBLrel. 25, Last annotation update)
smembrane protein.
03756.
anacearum (Pseudomonas solanacearum).
plasmid.
oteobacteria; Betaproteobacteria; Burkholderiales;
eae; Ralstonia.
5;

[N.A.
0;
879; PubMed=11823852;
Genin S., Artiguenave F., Gouzy J., Mangenot S.,
lault A., Brottier P., Camus J.C., Cattolico L.,
Choisne N., Claudel-Renard C., Cunnac S., Demange N.,
vie M., Moisan A., Robert C., Saurin W., Schiex T.,
hebault P., Whalen M., Wincker P., Levy M.,
., Boucher C.A.;
nce of the plant pathogen Ralstonia solanacearum.";
07-502(2002).
9; CAD17762.1; -.
1; C:extrachromosomal DNA; IEA.
4; F:ATP binding; IEA.
0; F:nucleoside-diphosphate kinase activity; IEA.
1; P:CTP biosynthesis; IEA.
3; P:GTP biosynthesis; IEA.
8; P:UTP biosynthesis; IEA.
001564; NDK.
469; NDP KINASES; 1.
lete proteome.
5 AA; 47048 MW; CCD859D9C54DDB5A CRC64;

arity   2.8%; Score 8; DB 16; Length 435;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

VLAL 224
     |||
VLAL 365

PRELIMINARY;   PRT;   465 AA.
TrEMBLrel. 20, Created)
TrEMBLrel. 20, Last sequence update)
TrEMBLrel. 25, Last annotation update)
brane protein (Putative permease, major facilitator

GN   YEGB OR YPO2850 OR Y1383.
OS   Yersinia pestis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Yersinia.
OX   NCBI_TaxID=632;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CO-92 / Biovar Orientalis;
RX   MEDLINE=21470413; PubMed=11586360;
RA   Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA   Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,
RA   Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA   Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA   Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA   Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA   Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT   "Genome sequence of Yersinia pestis, the causative agent of plag
RL   Nature 413:523-527(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KIM5 / Biovar Mediaevalis;
RX   MEDLINE=22137863; PubMed=1214230;
RA   Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Lis
RA   Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA   Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA   Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner
RA   Perry R.D.;
RT   "Genome sequence of Yersinia pestis KIM.";
RL   J. Bacteriol. 184:4601-4611(2002).
DR   EMBL; AJ414154; CAC92102.1; -.
DR   EMBL; AE013741; AAM84955.1; -.
DR   PIR; AC0347; AC0347.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0015520; F:tetracycline:hydrogen antiporter activity; IEA
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0015904; P:tetracycline transport; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR001411; TCR_TetB.
DR   Pfam; PF00083; sugar_tr; 1.
DR   PRINTS; PR01036; TCRTETB.
DR   PROSITE; PS50850; MFS; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE 465 AA; 50176 MW; 0CC273F10B3F5ED CRC64;

Query Match   2.8%; Score 8; DB 16; Length 465;
Best Local Similarity   100.0%; Pred. No. 60;
Matches   8; Conservative   0; Mismatches   0; Indels   0; C

Qy   72 LLLAVVSL 79
          ||||||||
Db   335 LLLAVVSL 342

RESULT 59
Q9RRI8
ID   Q9RRI8   PRELIMINARY;   PRT;   471 AA.
AC   Q9RRI8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Transport protein, putative.
GN   DR2502.
OS   Deinococcus radiodurans.
OC   Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC   Deinococcaceae; Deinococcus.
OX   NCBI_TaxID=1299;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX   MEDLINE=20036896; PubMed=10567266;
RA   White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,

```
aft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
in H., Jiang L., Pamphile W., Crosby M., Shen M.,
., Lam P., McDonald L.J., Utterback T., Zalewski C.,
Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,

ce of the radioresistant bacterium Deinococcus
.";
71-1577(1999).
.; AAF12043.1; -.
75267.
-.
ome.
    AA; 47974 MW; 96B2BBF6E445D27 CRC64;

ity      2.8%;  Score 8;  DB 16;  Length 471;
         100.0%;  Pred. No. 61;
nservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

LAL  67
     |||
LAL  373
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC023417; AAH23417.1; -.
DR   MGD; MGI:1924832; 5730589L02Rik.
DR   InterPro; IPR004299; MBOAT_fam.
DR   Pfam; PF03062; MBOAT; 1.
SQ   SEQUENCE 473 AA; 53382 MW; DAA1FE0DA78013EA CRC64;

Query Match          2.8%;  Score 8;  DB 11;  Length 473;
Best Local Similarity  100.0%;  Pred. No. 61;
Matches      8;  Conservative    0;  Mismatches    0;  Indels    0;  G

Qy    65  LALACLGL  72
          ||||||||
Db   436  LALACLGL  443
```

```
RESULT 62
Q9CY76
ID   Q9CY76       PRELIMINARY;      PRT;   473 AA.
AC   Q9CY76;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   5730589L02Rik protein.
GN   5730589L02RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilmin
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK019981; BAB31950.1; -.
DR   MGD; MGI:1924832; 5730589L02Rik.
DR   InterPro; IPR004299; MBOAT_fam.
DR   Pfam; PF03062; MBOAT; 1.
SQ   SEQUENCE 473 AA; 53504 MW; CE6F8E93C3D01C4F CRC64;

Query Match          2.8%;  Score 8;  DB 11;  Length 473;
Best Local Similarity  100.0%;  Pred. No. 61;
Matches      8;  Conservative    0;  Mismatches    0;  Indels    0;  G

Qy    65  LALACLGL  72
          ||||||||
Db   436  LALACLGL  443
```

```
ELIMINARY;      PRT;   472 AA.

rEMBLrel. 19, Created)
rEMBLrel. 19, Last sequence update)
rEMBLrel. 22, Last annotation update)
rotein FLJ31346.
Human).
azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
eria; Primates; Catarrhini; Hominidae; Homo;
6;

N.A.
mazaki M., Watanabe K., Kumagai A., Itakura S.,
ujimori Y., Komiyama M., Sugiyama T., Irie R.,
o H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
Saito K., Nishikawa T., Kimura K., Yamashita H.,
amura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
anehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
uzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,

DNA sequencing project.";
-2001) to the EMBL/GenBank/DDBJ databases.
3; BAB71043.1; -.
004299; MBOAT_fam.
MBOAT; 1.
rotein.
    AA; 52774 MW; EA72199B043F9EBD CRC64;

ity      2.8%;  Score 8;  DB 4;  Length 472;
         100.0%;  Pred. No. 61;
nservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

LAL  67
     |||
LAL  447
```

```
RELIMINARY;      PRT;   473 AA.

rEMBLrel. 21, Created)
rEMBLrel. 21, Last sequence update)
rEMBLrel. 22, Last annotation update)
30589L02 gene.
(Mouse).
```

```
Query Match          2.8%;  Score 8;  DB 11;  Length 473;
Best Local Similarity  100.0%;  Pred. No. 61;
Matches      8;  Conservative    0;  Mismatches    0;  Indels    0;  G

Qy    65  LALACLGL  72
          ||||||||
Db   436  LALACLGL  443
```

RESULT 65
ID   Q9LIW0   PRELIMINARY;   PRT;   522 AA.
AC   Q9LIW0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Similar to an Arabidopsis thialiana chromosome BAC genomic
DE   sequence.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hsing Y.C., Chow T., Chen C., Wu H., Chu M., Chao Y., Liu S.;
RT   "Oryza sativa PAC P0699E04 genomics sequence, complete sequence.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP001111; BAA90509.1; -.
DR   Gramene; Q9LIW0; -.
SQ   SEQUENCE   522 AA;   54697 MW; 21C6BAD2441B56BF CRC64;

Query Match   2.8%; Score 8; DB 10; Length 522;
Best Local Similarity   100.0%; Pred.No. 67;
Matches   8; Conservative   0; Mismatches   0; Indels   0; (

Qy   43 RRRGRRGE 50
             ||||||||
Db   415 RRRGRRGE 422

RESULT 66
ID   Q7WZ71   PRELIMINARY;   PRT;   535 AA.
AC   Q7WZ71;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Putative mannosyltransferase.
GN   DBV20.
OS   Nonomuraea sp. ATCC 39727.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptosporangineae; Streptosporangiaceae; Nonomuraea.
OX   NCBI_TaxID=93944;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 39727;
RA   Sosio M., Stinchi S., Beltrametti F., Lazzarini A., Donadio S.;
RT   "The gene cluster for the biosynthesis of the glycopeptide antib
RT   A40926 by Nonomuraea sp.";
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ561198; CAD91215.1; -.
KW   Acyltransferase; Glycosyltransferase; Monooxygenase; Transferase
SQ   SEQUENCE   535 AA;   57418 MW; 3C9059338B3308AC CRC64;

Query Match   2.8%; Score 8; DB 2; Length 535;
Best Local Similarity   100.0%; Pred.No. 68;
Matches   8; Conservative   0; Mismatches   0; Indels   0; (

Qy   122 RLVRPRRS 129
             ||||||||
Db   480 RLVRPRRS 487

RESULT 67
ID   Q8PMH8   PRELIMINARY;   PRT;   537 AA.
AC   Q8PMH8;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Oligopeptide transporter.

PRELIMINARY;   PRT;   473 AA.

(Mouse).
TrEMBLrel. 23, Created)
TrEMBLrel. 23, Last sequence update)
TrEMBLrel. 24, Last annotation update)

OR MBB1.
tazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
heria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
090;
N.A.
ayama I., Fujino M.A.;
nd characterization of novel human and mouse genes, which
in the digestive tract.";
B-2001) to the EMBL/GenBank/DDBJ databases.
0; BAC53808.1; -.
832; 5730589L02Rik.
004299; MBOAT_fam.
; MBOAT; 1.
3 AA; 53435 MW; CFB4FB0DB2951C4F CRC64;

arity   2.8%; Score 8; DB 11; Length 473;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

CLGL 72
||||
CLGL 443

PRELIMINARY;   PRT;   496 AA.

TrEMBLrel. 23, Created)
TrEMBLrel. 23, Last sequence update)
TrEMBLrel. 25, Last annotation update)
protein.
(japonica cultivar-group).
ridiplantae; Streptophyta; Embryophyta; Tracheophyta;
; Magnoliophyta; Liliopsida; Poales; Poaceae;
; Oryzeae; Oryza.
947;
N.A.
pponbare;
tsumoto T., Yamamoto K.;
nipponbare(GA3) genomic DNA, chromosome 7, BAC
E01.";
L-2001) to the EMBL/GenBank/DDBJ databases.
29; BAC45051.1; -.
6; F:nucleic acid binding; IEA.
000571; Znf CCCH.
; zf-CCCH; 3.
56; ZnF_C3H1; 3.
6 AA; 51701 MW; 3AFBA80FC0AFA115 CRC64;

arity   100.0%; Score 8; DB 10; Length 496;
Conservative   0; Mismatches   0; Indels   0; Gaps   0;

VDGV 221
|||||
VDGV 208

DR   GO; GO:0019031; C:viral envelope; IEA.
DR   InterPro; IPR008980; Capsid_hemag.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Glycoprotein; Hemagglutinin.
FT   NON_TER   564   564
SQ   SEQUENCE   564 AA; 63562 MW; B317179A7F3E6F98 CRC64;

Query Match   2.8%; Score 8; DB 12; Length 564;
Best Local Similarity   100.0%; Pred. No. 71;
Matches   8; Conservative   0; Mismatches   0; Indels   0; G

Qy   72 LLLAVVSL 79
            ||||||||
Db   6 LLLAVVSL 13

RESULT 69
Q8JN92
ID   Q8JN92   PRELIMINARY;   PRT;   568 AA.
AC   Q8JN92;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hemagglutinin H5.
GN   H5.
OS   Influenza A virus (A/Goose/Hong Kong/3014.5/2000(H5N1)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza A viruses; Influenzavirus A.
OX   NCBI_TaxID=186167;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/Goose/Hong Kong/3014.5/2000;
RX   MEDLINE=22016166; PubMed=12021367;
RA   Tumpey T.M., Suarez D.L., Perkins L.E.L., Senne D.A., Lee J.G.,
RA   Lee Y.J., Mo I.P., Sung H.W., Swayne D.E.;
RT   "Characterization of a Highly Pathogenic H5N1 Avian Influenza A V
RT   isolated from Duck Meat.";
RL   J. Virol. 76:6344-6355(2002).
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRU
CC   CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CH
CC   (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR   EMBL; AY075030; AAL75843.1; -.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   InterPro; IPR008980; Capsid_hemag.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Glycoprotein; Hemagglutinin.
SQ   SEQUENCE   568 AA; 64244 MW; E0D741A75C8E76FC CRC64;

Query Match   2.8%; Score 8; DB 12; Length 568;
Best Local Similarity   100.0%; Pred. No. 72;
Matches   8; Conservative   0; Mismatches   0; Indels   0; G

Qy   72 LLLAVVSL 79
            ||||||||
Db   6 LLLAVVSL 13

RESULT 70
Q8QPL0
ID   Q8QPL0   PRELIMINARY;   PRT;   568 AA.
AC   Q8QPL0;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hemagglutinin (Fragment).

0.
onopodis (pv. citri).
eobacteria; Gammaproteobacteria; Xanthomonadales;
ae; Xanthomonas;
29;

N.A.
TCC 13902 / XV 101;
45; PubMed=12024217;
., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
erreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
F., Franco M.C., Greggio C.C., Gruber A.,
., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
achado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
, Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
santos M., Truffi D., Tsai S.M., White F.F.,
Kitajima J.P.;
the genomes of two Xanthomonas pathogens with differing
ties.";
-463(2002).
; AAM36320.1; -.
; C:membrane; IEA.
.; F:transporter activity; IEA.
.; P:oligopeptide transport; IEA.
00109; PTR2.
PTR2; 1.
ome.
AA; 58369 MW; 7844C0C0FE1E8670 CRC64;

2.8%; Score 8; DB 16; Length 537;
rity 100.0%; Pred. No. 68;
nservative   0; Mismatches   0; Indels   0; Gaps   0;

VVS 78
 |||
VVS 263

ELIMINARY;   PRT;   564 AA.

rEMBLrel. 10, Created)
rEMBLrel. 10, Last sequence update)
rEMBLrel. 25, Last annotation update)
H5 (Fragment).
.A.
negative-strand viruses; Orthomyxoviridae;
ruses.
911;

N.A.
Potsdam/2216-4/84;
02; PubMed=9882316;
, Zhou N., Kawaoka Y., Webster R.;
lycoproteins of H5 influenza viruses isolated from
ns, and wild aquatic birds have distinguishable

146-1155(1999).
HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
TORS AND FOR INITIATING INFECTION (BY SIMILARITY).
OMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
A2) LINKED BY A DISULFIDE BOND (BY SIMILARITY)
: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
; AAD13573.1; -.
1HTM.

RT host specificities.";
RL Nature 417:459-463(2002).
DR EMBL; AE012240; AAM40703.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006857; P:oligopeptide transport; IEA.
DR InterPro; IPR000109; PTR2.
DR Pfam; PF00854; PTR2; 1.
DR PROSITE; PS01022; PTR2_1; 1.
KW Complete proteome.
SQ SEQUENCE 620 AA; 67314 MW; E9904BFF039B6AEC CRC64;

Query Match 2.8%; Score 8; DB 16; Length 620;
Best Local Similarity 100.0%; Pred. No.78;
Matches 8; Conservative 0; Mismatches 0; Indels 0;

Qy 65 LALACLGL 72
   ||||||||
Db 349 LALACLGL 356

RESULT 72
Q9N8H2
ID Q9N8H2 PRELIMINARY; PRT; 656 AA.
AC Q9N8H2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN TB927.1.3840.
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanos
OX NCBI_TaxID=5691;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TREU927;
RA Hall N., Berriman M., Lennard N.J., Harris B.R., Gerrard C.S.,
RA Atkin R.J., Barron A.J., Bart-Delabesse E.N., Bowman S.,
RA Bray-Allen S.P., Bringaud F., Clark L.N., Corton C.H., Cronin A.,
RA Davies R., Doggett J., Fraser A., Gruter E., Hall S., Harper D.A
RA Hertz-Fowler C., Kay M.P., Leech V., Mayes R., Price C., Quail M
RA Rabbinowitsch E., Rutherford K., Sasse J., Sharp S., Shownkeen R
RA Gull K., Barrell B.G., Melville S.E.;
RT "The sequence and analysis of the highly polymorphic chromosome
RT the African trypanosome, Trypanosoma brucei.";
RL Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
EMBL; AL929607; CAB95571.1; -.
KW Hypothetical protein.
SQ SEQUENCE 656 AA; 72138 MW; CBAC892D25937FAD CRC64;

Query Match 2.8%; Score 8; DB 5; Length 656;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 8; Conservative 0; Mismatches 0; Indels 0;

Qy 93 EEIVAEED 100
   ||||||||
Db 454 EEIVAEED 461

RESULT 73
Q9N8U8
ID Q9N8U8 PRELIMINARY; PRT; 1523 AA.
AC Q9N8U8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN TB927.1.1600.
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanos
OX NCBI_TaxID=5691;
RN [1]
RP SEQUENCE FROM N.A.

---

virus (A/Goose/Hong Kong/3014.8/2000(H5N1)).
A negative-strand viruses; Orthomyxoviridae;
viruses; Influenzavirus A.
6675;
N.A.
e/Hong Kong/3014.8/2000(H5N1);
832; PubMed=1878904;
is M., Kong K.F., Dyrting K.C., Ellis T.M., Sit T.,
hortridge K.F.;
za viruses isolated from Geese in Southeastern China:
Genetic Reassortment and Interspecies Transmission to

16-23(2002).
HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
PTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
Y: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
2; AAL31388.1; -.
1; C:viral envelope; IEA.
008980; Capsid_hemag.
001364; Hemagglutn.
; Hemagglutinin; 1.
29; HEMAGGLUTN12.
225; Hemagglutn; 1.
ein; Glycoprotein; Hemagglutinin.
06.
68 568
8 AA; 64281 MW; 0B0A4CFE034F1769 CRC64;

arity 2.8%; Score 8; DB 12; Length 568;
 100.0%; Pred. No. 72;
onservative 0; Mismatches 0; Indels 0; Gaps 0;

VVSL 79
||||
VVSL 13

PRELIMINARY; PRT; 620 AA.

TrEMBLrel. 22, Created)
TrEMBLrel. 22, Last sequence update)
TrEMBLrel. 24, Last annotation update)
transporter.
06.
ampestris (pv. campestris);
teobacteria; Gammaproteobacteria; Xanthomonadales;
eae; Xanthomonas.
0;
N.A.
3913 / NCPPB 528;
145; PubMed=i1024217;
R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
erreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
.F., Franco M.C., Greggio C.C., Gruber A.,
., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
Santos M., Truffi D., Tsai S.M., White F.F.,
Kitajima J.P.;
f the genomes of two Xanthomonas pathogens with differing

RP SEQUENCE FROM N.A.
RC STRAIN=L2 434B;
RA Wang L., Steenburg S.D., Zheng Y., Larsen S.H.;
RT "Gene identification of Chlamydia trachomatis by random DNA
RT sequencing.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF087333; AAD04107.1; -.
KW Hypothetical protein.
FT NON_TER   1   1
FT NON_TER   35   35
SQ SEQUENCE   35 AA;   3519 MW;   46686B72BAF28F7D CRC64;

Query Match   2.5%;   Score 7; DB 2;   Length 35;
Best Local Similarity   100.0%;   Pred. No. 57;
Matches   7; Conservative   0; Mismatches   0; Indels   0; G

Qy   174 NSSSPLR 180
          |||||||
Db   8 NSSSPLR 14

Search completed: April 7, 2004, 17:58:55
Job time : 50 secs

7;
iman M., Lennard N.J., Harris B.R., Gerrard C.S.,
arron A.J., Bart-Delabesse E.N., Bowman S.,
., Bringaud F., Clark L.N., Corton C.H., Cronin A.,
ggett J., Fraser A., Gruter E., Hall S., Harper D.A.,
., Kay M.P., Leech V., Mayes R., Price C., Quail M.A.,
E., Rutherford K., Sasse J., Sharp S., Shownkeen R.,
ell B.G., Melville S.E.;
and analysis of the highly polymorphic chromosome I of
rypanosome, Trypanosoma brucei.";
-2002) to the EMBL/GenBank/DDBJ databases.
1; CAB95435.1; -.
001656; UPF0024.
; UPF0024; 1.
protein.
23 AA;   168322 MW;   077BDC751CDD1E5A CRC64;

arity   2.8%;   Score 8; DB 5;   Length 1523;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

ALR 225
    |||
ALR 389

ELIMINARY;   PRT;   32 AA.

rEMBLrel. 17, Created)
rEMBLrel. 17, Last sequence update)
rEMBLrel. 19, Last annotation update)
ULIN.
Goat).
azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
eria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
nae; Capra.
5;
N.A.
illa F., Graziano M.;
phism of goat beta-lactoglobulin proximal promoter
-2001) to the EMBL/GenBank/DDBJ databases.
; CAC27455.1; -.
             POTENTIAL.
1   18
2   32
AA;   3372 MW;   0C56BD579B3DC190 CRC64;

rity   2.5%;   Score 7; DB 6;   Length 32;
nservative   0; Mismatches   0; Indels   0; Gaps   0;

AC 69
   ||
AC 14

ELIMINARY;   PRT;   35 AA.

rEMBLrel. 10, Created)
rEMBLrel. 10, Last sequence update)
rEMBLrel. 25, Last annotation update)
rotein (Fragment).
homatis.
mydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
;

GenCore version 5.1.6
opyright (c) 1993 - 2004 Compugen Ltd.

n search, using sw model

il 7, 2004, 17:47:28 ; Search time 60 Seconds
(without alignments)
1337.392 Million cell updates/sec

09-245-198A-4

SLLDFEISARRLPLPRSLG.........PWAHLKAAPFLTYFGLFQVH 284

GO
op 60.0 , Gapext 60.0

6107 seqs, 282547505 residues

s satisfying chosen parameters: 1586107

rth: 0
rth: 2000000000

sting first 100 summaries

Geneseq.29Jan04:*
geneseqp1980s:*
geneseqp1990s:*
geneseqp2000s:*
geneseqp2001s:*
geneseqp2002s:*
geneseqp2003as:*
geneseqp2003bs:*
geneseqp2004s:*

the number of results predicted by chance to have a
than or equal to the score of the result being printed,
d by analysis of the total score distribution.

SUMMARIES

ry ch	Length	DB	ID	Description					
.0	284	2	AAW47525	Aaw47525 Homo sapi					
.7	249	3	AAY09369	Aay09369 Human tum					
.7	249	3	AAY95338	Aay95338 Human PRO					
.7	249	3	AAB07526	Aab07526 Amino aci					
.7	249	5	AAU86129	Aau86129 Human PRO					
.7	249	6	ABR42315	Abr42315 Human TWE					
.7	249	7	ADC35206	Adc35206 Human TNF					
.9	249	5	AAW29745	Aaw29745 TNF relat					
.9	249	4	AAE00891	Aae00891 Human TRE					
.9	273	4	AAU03499	Aau03499 TWEAK ext					
.4	146	4	AAE00895	Aae00895 Human TRE					
.4	189	2	AAW29746	Aaw29746 TNF relat					
.4	189	4	AAE00892	Aae00892 Human UL4					
.7	208	2	AAW93590	Aaw93590 Human TNR					
.2	211	2	AAW93591	Aaw93591 Mouse TNR					
.3	225	2	AAW47524	Aaw47524 Mus muscu					
.3	225	4	AAB07527	Aab07527 Amino aci					
.3	249	7	ADC97712	Adc97712 Murine FL					
.2	58	3	AAG01265	Aag01265 Human sec					
.2	58	4	AAG01266	Aag01266 Human sec					
.2	365	6	ABR41235	Abr41235 Human DIT					
.2	748	2	AAY14906	Aay14906 Extended					
.2	749	5	ABB73512	Abb73512 M vaccae					
.8	54	4	AAU51863	Aau51863 Propionib					
.8	54	6	ABM48382	Abm48382 Propionib					
26	8	2.8	55	4	AAM20858	Aam20858	Pej		
27	8	2.8	55	4	AAB42637	Abb42637	Pej		
28	8	2.8	55	4	AAM36451	Aam36451	Pej		
29	8	2.8	55	4	ABB25983	Abb25983	Pri		
30	8	2.8	55	4	AAM76342	Aam76342	Hur		
31	8	2.8	55	4	AAM63528	Aam63528	Hur		
32	8	2.8	55	4	ABG58050	Abg58050	Hur		
33	8	2.8	55	5	ABG45635	Abg45635	Hur		
34	8	2.8	65	4	AAM21621	Aam21621	Pej		
35	8	2.8	65	4	ABB43981	Abb43981	Pej		
36	8	2.8	65	4	AAM37923	Aam37923	Pej		
37	8	2.8	65	4	ABB26890	Abb26890	Pri		
38	8	2.8	65	4	AAM77706	Aam77706	Hur		
39	8	2.8	65	4	AAM64984	Aam64984	Hur		
40	8	2.8	65	4	ABG59361	Abg59361	Hur		
41	8	2.8	65	5	ABG46737	Abg46737	Hur		
42	8	2.8	69	4	ABG03723	Abg03723	No'		
43	8	2.8	71	4	ABG03663	Abg03663	No'		
44	8	2.8	84	4	AAU61282	Aau61282	Pri		
45	8	2.8	84	6	ABM57801	Abm57801	Pri		
46	8	2.8	110	4	ABG20259	Abg20259	No'		
47	8	2.8	117	4	AAO08094	Aao08094	Hur		
48	8	2.8	184	4	ABB67712	Abb67712	Dr:		
49	8	2.8	190	5	ABP28041	Abp28041	St:		
50	8	2.8	198	4	AAM23684	Aam23684	Hur		
51	8	2.8	222	5	ABP41674	Abp41674	Hur		
52	8	2.8	286	4	ABG16279	Abg16279	No'		
53	8	2.8	307	4	ABG01186	Abg01186	No'		
54	8	2.8	342	7	ADB79952	Adb79952	Ral		
55	8	2.8	370	4	ABG13391	Abg13391	No'		
56	8	2.8	370	4	ABG05012	Abg05012	No'		
57	8	2.8	370	4	ABG18115	Abg18115	No'		
58	8	2.8	372	6	ABU36537	Abu36537	Pri		
59	8	2.8	424	4	ABG15613	Abg15613	No'		
60	8	2.8	430	4	AAM23752	Aam23752	Hur		
61	8	2.8	431	7	ADC08203	Adc08203	Ri		
62	8	2.8	431	7	ADC64562	Adc64562	Sy		
63	8	2.8	454	7	ABB57908	Abb57908	Dr		
64	8	2.8	472	6	ADA54710	Ada54710	Hur		
65	8	2.8	586	4	ABG20260	Abg20260	No'		
66	8	2.8	586	7	ADD46285	Add46285	Hur		
67	8	2.8	586	7	ADE62980	Ade62980	Hur		
68	8	2.8	592	4	AAU32148	Aau32148	No'		
69	8	2.8	603	4	ABG03722	Abg03722	No'		
70	8	2.8	617	7	ADE79006	Ade79006	Hur		
71	8	2.8	633	4	ABG20261	Abg20261	No'		
72	8	2.8	799	4	ABG28291	Abg28291	No'		
73	8	2.8	842	7	ADE08475	Ade08475	No'		
74	8	2.8	1032	6	ABU16705	Abu16705	Prc		
75	8	2.8	1033	6	ADA33851	Ada33851	Ac:		
76	7	2.5	12	6	AAR66780	Aar66780	Ce:		
77	7	2.5	16	6	ABP82261	Abp82261	G.]		
78	7	2.5	17	5	ABB09088	Abb09088	Hur		
79	7	2.5	18	5	ABG62026	Abg62026	Hur		
80	7	2.5	19	6	ADA50609	Ada50609	HC		
81	7	2.5	20	7	ADC99453	Adc99453	Car		
82	7	2.5	23	5	ABG62030	Abg62030	Hur		
83	7	2.5	24	2	AAY34189	Aay34189	Hur		
84	7	2.5	25	3	AAB18622	Aab18622	Ant		
85	7	2.5	44	2	AAW89072	Aaw89072	Po:		
86	7	2.5	44	4	AAE01449	Aae01449	Hur		
87	7	2.5	44	4	AAE01489	Aae01489	Hur		
88	7	2.5	44	5	ABB51243	Abb51243	Hur		
89	7	2.5	44	5	ABG63884	Abg63884	Hun		
90	7	2.5	44	5	ABG63885	Abg63885	Hur		
91	7	2.5	44	5	ABO45500	Abo45500	Hur		
92	7	2.5	44	7	ABO26980	Abo26980	Prc		
93	7	2.5	46	4	ABG02779	Abg02779	No'		
94	7	2.5	55	3	AAB40432	Aab40432	Hur		
95	7	2.5	55	4	AAU63542	Aau63542	Hur		
96	7	2.5	55	6	ABM60061	Abm60061	Prc		
97	7	2.5	58	5	ABP00382	Abp00382	Hur		
98	7	2.5	65	2	AAY12852	Aay12852	Hur		

```
2.5    66   7   ADC97100        Adc97100 E. faeciu
2.5    69   4   ABB65500        Abb65500 Drosophil

              ALIGNMENTS
```

dardard; protein; 284 AA.

(first entry)

tumour necrosis factor related ligand (TRELL).

r necrosis factor related ligand; tnf; treatment; cancer;
isease; immune system; stimulation; suppression;
ion.

97WO-US013945.

96US-0023541P.
96US-0028515P.
97US-0040820P.

EN INC.
GENEVA FACULTY MEDICINE.

he Y, Browning JL;

5619/13.
8600.

sis factor related ligand - useful for, e.g. treating cancer,
disease and immune responses to tissue grafts.

ge 50-51; English.

```
Db    1 MSLLDFEISARRLPLPRSLGSRDGGAVRQAQPPAPMAARRSQRRRGRRGEPGTAL
Qy   61 LGLGLALACLGLLLAVVSLGSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQD
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LGLGLALACLGLLLAVVSLGSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQD
Qy  121 NRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINS
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 NRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINS
Qy  181 YNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYLKLDLLVDGVLALRCLEEFSATAA
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 YNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYLKLDLLVDGVLALRCLEEFSATAA
Qy  241 QLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 284
           ||||||||||||||||||||||||||||||||||||||||||||
Db  241 QLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 284


RESULT 2
AAY09369
ID   AAY09369 standard; protein; 249 AA.
XX
AC   AAY09369;
XX
DT   15-JUL-1999  (first entry)
XX
DE   Human tumour necrosis factor Apo-3 ligand protein sequence.
XX
KW   Human; tumour necrosis factor; Apo-3 ligand; lymphotoxin; apopto
KW   NF-kappaB-dependent transcription; JNK/SAPK-dependent response;
XX
OS   Homo sapiens.
XX
XX
PN   WO9919490-A1.
XX
PD   22-APR-1999.
XX
PF   09-OCT-1998;   98WO-US021407.
XX
PR   10-OCT-1997;   97US-0062037P.
PR   17-DEC-1997;   97US-0069862P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Ashkenazi AJ, Marsters SA,  Pitti R;
XX
DR   WPI; 1999-287982/24.
DR   N-PSDB; AAX56000.
XX
PT   New human Apo3- ligand (a tumor necrosis factor) homologue.
XX
PS   Claim 1; Fig 1; 74pp; English.
XX
CC   The present sequence represents a human tumour necrosis factor (
CC   lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has
CC   cytostatic activity. Apo-3 ligand can be used to induce apoptosi
CC   mammalian cancer cells, to induce NF-kappaB-dependent transcript
CC   to induce JNK/SAPK-dependent responses in mammalian cells
XX
SQ   Sequence 249 AA;

Query Match    87.7%;   Score 249;  DB 2;  Length 249;
Best Local Similarity 100.0%;  Pred. No. 2.6e-220;
Matches 249;  Conservative  0;  Mismatches  0;  Indels  0;  (

Qy   36 MAARRSQRRRGRRGEPGTALLVPLALGLGLALACLGLLLAVVSLGSRASLSAQEP
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MAARRSQRRRGRRGEPGTALLVPLALGLGLALACLGLLLAVVSLGSRASLSAQEP
Qy   96 VAEEDQDPSELNPQTEESQDPAPFINRLVRPRRSAPKGRKTRARRAIAAHYEVHP
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VAEEDQDPSELNPQTEESQDPAPFINRLVRPRRSAPKGRKTRARRAIAAHYEVHP
```

AA;

```
larity 100.0%;  Score 284;  DB 2;  Length 284;
larity 100.0%;  Pred. No. 2e-252;
Conservative   0;  Mismatches  0;  Indels  0;  Gaps  0;

LDFEISARRLPLPRSLGSRDGGAVRQAQPPAPMAARRSQRRRGRRGEPGTALLVPLA 60
```

```
PR   26-JUL-1999;   99US-0145698P.
PR   15-SEP-1999;   99WO-US021090.
PR   15-SEP-1999;   99WO-US021547.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA;
PI   Napier MA, Pitti RM, Wood WI;
XX
DR   WPI; 2000-442668/38.
DR   N-PSDB; AAA49717.
XX
PT   Novel composition to inhibit neoplastic cell growth or for treati
PT   in mammal comprises polypeptides PRO179, PRO207, PRO219,
PT   PRO224, PRO328, PRO301, PRO526, PRO362, PRO509 or PRO866.
XX
PS   Claim 19; Fig 4; Fig 4; 172pp; English.
XX
CC   The present sequence is that of human antitumour protein PRO207,
CC   deduced from a foetal kidney cDNA clone (see AAA49717). PRO207 sh
CC   amino acid sequence identity to tumour necrosis factor family mem
CC   especially human lymphotoxin-beta (23.4%) and human CD40 ligand (
CC   Mol.wt. is 27,216. A claimed method for inhibiting the growth of
CC   cell comprises exposing the tumor cell to PRO179, PRO207, PRO320,
CC   PRO221, PRO224, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or
CC   (see AAY95337-49), their agonists or chimeric polypeptides incorp
CC   them. The tumour can be selected a cancer selected from breast, ova
CC   renal, colorectal, uterine, prostate, lung, bladder and central n
CC   system cancer, melanoma and leukaemia. Methods for the recombinan
CC   expression of the antitumour proteins are also provided
XX
SQ   Sequence 249 AA;

Query Match             87.7%;   Score 249;  DB 3;  Length 249;
Best Local Similarity  100.0%;   Pred. No. 2.6e-220;
Matches  249;   Conservative   0;   Mismatches   0;   Indels   0;   G

Qy    36  MAARRSQRRRGRRGEPGTALLVPLAIGLGLLAVVSLGSRASLSAQEPA
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  MAARRSQRRRGRRGEPGTALLVPLAIGLGLLAVVSLGSRASLSAQEPA

Qy    96  VAEEDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPR
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  VAEEDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPR

Qy   156  GAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAVY
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  GAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAVY

Qy   216  LLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKA
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  LLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKA

Qy   276  TYFGLFQVH  284
              |||||||||
Db   241  TYFGLFQVH  249

RESULT 4
AAB07526
ID   AAB07526 standard; protein; 249 AA.
XX
AC   AAB07526;
XX
DT   20-OCT-2000 (first entry)
XX
DE   Amino acid sequence of a soluble recombinant human TWEAK protein.
XX
KW   TWEAK protein; immunological disorder; immune response; inflammat
KW   TWEAK blocking agent; autoimmune disease; organ transplant reject
KW   Graft-versus-Host disease; GVHD; lymphoid cell malignancy; shock;
XX
OS   Homo sapiens.
```

Left portion:

```
VDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYLKLD  215
|||||||||||||||||||||||||||||||||||||||||||||||||
VDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYLKLD  180

VLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFL  275
|||||||||||||||||||||||||||||||||||||||||||||||||
VLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFL  240

FQVH  284
||||
FQVH  249

lard; protein; 249 AA.

(first entry)

ntitumour protein.

antitumour; tumour; therapy; cytostatic; breast cancer;
; renal cancer; colorectal cancer; uterine cancer;
r; lung cancer; bladder cancer;
s system cancer; melanoma; leukaemia; neoplasm.

Location/Qualifiers
1..40
/label= Signal_peptide
10..14
/note= "amidation"
24..35
/note= "prokaryotic membrane lipoprotein lipid"
27..33
/note= "N-myristoylation"
29..35
/note= "N-myristoylation"
36..42
/note= "N-myristoylation"
41..249
/label= PRO207
45..51
/note= "N-myristoylation"
97..101
/note= "amidation"
118..124
/note= "N-myristoylation"
121..127
/note= "N-myristoylation"
125..131
/note= "N-myristoylation"
128..134
/note= "N-myristoylation"
139..143
/note= "Asn is N-glycosylated"
2.

99WO-US028565.

98US-0113296P.
99WO-US005028.
99US-0130232P.
99US-0131445P.
99US-0134287P.
99US-0144758P.
```

KW   neuroprotective.
XX
OS   Homo sapiens.
XX
PN   WO20153486-A1.
XX
PD   26-JUL-2001.
XX
PF   11-FEB-2000; 2000WO-US003565.
XX
PR   08-MAR-1999;   99WO-US005028.
PR   11-MAR-1999;   99US-0123972P.
PR   11-MAY-1999;   99US-0133459P.
PR   02-JUN-1999;   99WO-US012252.
PR   22-JUN-1999;   99US-0140650P.
PR   22-JUN-1999;   99US-0140653P.
PR   20-JUL-1999;   99US-0144758P.
PR   26-JUL-1999;   99US-0145698P.
PR   28-JUL-1999;   99US-0146222P.
PR   17-AUG-1999;   99US-0149395P.
PR   31-AUG-1999;   99US-0151689P.
PR   01-SEP-1999;   99WO-US020111.
PR   15-SEP-1999;   99WO-US021090.
PR   30-NOV-1999;   99WO-US028313.
PR   01-DEC-1999;   99WO-US028301.
PR   01-DEC-1999;   99WO-US028634.
PR   05-JAN-2000;   2000WO-US000219.
XX
PA   (GETH) GENENTECH INC.
XX
PI   Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL,   Hillan KJ;
PI   Marsters SA,   Pan J, Pitti RM, Roy MA,   Smith V,   Stone DM;
PI   Watanabe CK,   Wood WI;
XX
DR   WPI; 2002-205567/26.
DR   N-PSDB; ABK40255.
XX
PT   Thirty five nucleic acids encoding PRO polypeptides, useful for
PT   benign or malignant tumors, leukemias and lymphoid malignancies,
PT   inflammatory, angiogenic and immunologic disorders.
XX
PS   Claim 61; Fig 4; 302pp; English.
XX
CC   The present invention relates to the isolation of novel human PR
CC   polypeptides and the polynucleotide sequences encoding them. The
CC   polypeptides, agonists, antagonists or anti-PRO antibodies are u
CC   treating benign or malignant tumours (e.g. renal, kidney, bladde
CC   breast, etc), leukaemias and lymphoid malignancies, other disord
CC   as neuronal, glial, astrocytal, hypothalamic, glandular, macroph
CC   stromal and blastocoelic disorders, inflammatory, immune and ang
CC   disorders. The polynucleotide sequences are also useful in gene
CC   AAU86128-AAU86162 represent the human PRO polypeptides of the in
XX
SQ   Sequence 249 AA.

Query Match   87.7%; Score 249; DB 5; Length 249;
Best Local Similarity 100.0%; Pred. No. 2.6e-220;
Matches 249; Conservative 0; Mismatches 0; Indels   0; (

Qy   36 MAARRSQRRRGRRGEPGTALLVPLALGLGLALACLGLLLAVVSLGSRASLSAQEP
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MAARRSQRRRGRRGEPGTALLVPLALGLGLALACLGLLLAVVSLGSRASLSAQEP

Qy   96 VAEEDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHP
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VAEEDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHP

Qy   156 GAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAV
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAV

Qy   216 LLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLK
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||

(left column, rotated fragments)

A1.

2000WO-US001044.

99US-0116168P.

EN INC.

6036/41.

nd treating immune responses using modulators, especially
of TWEAK, TWEAK receptors and TWEAK ligands, useful for
. inflammation and graft versus host disease.

Fig 1; 45pp; English.

sequence represents a TWEAK protein. The specification
method for preventing or treating an immunological disorder
iting an immune response in an animal. The method comprises
g a TWEAK blocking agent. The method may be used for
nd treating immune disorders associated with inappropriate
nd/or activity of TWEAK. These disorders include autoimmune
ute and chronic inflammation, organ transplant rejection,
-Host disease (GVHD), lymphoid cell malignancies, septic and
of shock, loss of immune responsiveness (as seen in human
ency virus (HIV) infections) and failure of the immune
tumour growth

AA;

larity   87.7%;   Score 249; DB 3; Length 249;
Conservative 0; Mismatches 0; Indels   0; Gaps   0;

RRSQRRRGRRGEPGTALLVPLALGLGLALACLGLLLAVVSLGSRASLSAQEPAQEEL 95
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
RRSQRRRGRRGEPGTALLVPLALGLGLALACLGLLLAVVSLGSRASLSAQEPAQEEL 60

EDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQD 155
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
EDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQD 120

AGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYLKLD 215
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYLKLD 180

DGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFL 275
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFL 240

GLFQVH 284
||||||
GLFQVH 249

ndard; protein; 249 AA.

(first entry)

polypeptide.

benign tumour; malignant tumour; lymphoid malignancy;
auronal disorder; stromal disorder; blastocoelic disorder;
disorder; immune disorder; angiogenic disorder; cytostatic;

```
Qy   216 LLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLK
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 LLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLK

Qy   276 TYFGLFQVH 284
             ||||||||
Db   241 TYFGLFQVH 249

RESULT 7
ADC35206
ID   ADC35206 standard; protein; 249 AA.
XX
AC   ADC35206;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Human TNF ligand family member #12.
XX
KW   human; tumour necrosis factor; TNF ligand; endokine alpha;
KW   excessive bone resorption disorder; osteoporosis; Paget's disease
KW   arterial calcification.
XX
OS   Homo sapiens.
XX
XX
PN   US2003100074-A1.
XX
PD   29-MAY-2003.
XX
PF   15-AUG-2002; 2002US-00218547.
XX
PR   16-AUG-2001; 2001US-0312542P.
PR   30-OCT-2001; 2001US-0330761P.
XX
PA   (YUGG/) YU G.
PA   (NIJJ/) NI J.
PA   (ROSE/) ROSEN C A.
PA   (NARD/) NARDELLI B.
XX
PI   Yu G, Ni J, Rosen CA, Nardelli B;
XX
DR   WPI; 2003-696072/66.
DR   N-PSDB; ADC35205.
XX
PT   New Endokine alpha gene useful for preparing a composition for tr
PT   disease associated with excessive or insufficient bone resorptior
PT   osteoporosis, Paget's disease or arterial calcification.
XX
PS   Disclosure; SEQ ID NO 24; 145pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule encodi
CC   tumour necrosis factor family ligand. A composition comprising th
CC   isolated antibody or its fragment is used for treating an individ
CC   need of decreased level of endokine alpha activity. The endokine
CC   polypeptide present in a heterotrimeric complex is used for treat
CC   individual having a disorder associated with excessive bone resor
CC   e.g. osteoporosis, Paget's disease or arterial calcification. Tre
CC   individual having a disorder associated with insufficient bone re
CC   comprises administering an endokine alpha antagonist, which is th
CC   antibody that binds specifically to endokine alpha polypeptide. T
CC   present sequence represents the amino acid sequence of a tumour n
CC   factor family ligand.
XX
SQ   Sequence 249 AA;

Query Match            87.7%;  Score 249;  DB 7;  Length 249;
Best Local Similarity  100.0%;  Pred. No. 2.6e-220;
Matches  249;  Conservative  0;  Mismatches  0;  Indels  0;  G

Qy    36 MAARRSQRRRGRRGREPGTALLVPLAIGLGLLAVVSLGSRASLSAQEPA
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MAARRSQRRRGRRGREPGTALLVPLAIGLGLLAVVSLGSRASLSAQEPA
```

```
GVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFL 240

LFQVH 284
 |||||
LFQVH 249

dard; protein; 249 AA.

(first entry)

rotein.

tumour necrosis factor; ligand; cytostatic;
or; osteopathic.

A2.

2002WO-US023782.

2001US-0307838P.

GENOME SCI INC.

Rosen CA;

659/40.
901.

timeric complex having a first polypeptide member of the
s factor (TNF) ligand family, and a second different member
family, useful for treating cancer, osteoporosis or an
sease.

age 368-369; 388pp; English.

equence is the protein sequence for human TWEAK protein. The
ates to compositions comprising heterotrimeric complexes of
is factor (TNF) ligand family members, and their use in the
evention and treatment of disease. In one embodiment, the
c complex comprises full-length or extracellular portions of
l-length or extracellular portions of other TNF ligand
s, preferably VEGI or VEGI-SV. The heterotrimeric complexes
ion are useful for treating an autoimmune disease, cancer or
and particularly for inhibiting cancer cell proliferation,
cell proliferation, or inducing apoptosis of T cells

AA;

arity    87.7%;  Score 249;  DB 6;  Length 249;
onservative  100.0%;  Pred. No. 2.6e-220;
             0;  Mismatches  0;  Indels  0;  Gaps  0;

RSQRRRGRRGREPGTALLVPLAIGLGLALACLGLLAVVSLGSRASLSAQEPAQEEL 95
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
RSQRRRGRRGREPGTALLVPLAIGLGLALACLGLLAVVSLGSRASLSAQEPAQEEL 60

DQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQD 155
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
DQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQD 120

3VDGTVSGWEARINSSSPLRYNRQIGEFIVTRAGLYLYLYCQVHFDEGKAVYLKLD 215
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
3VDGTVSGWEARINSSSPLRYNRQIGEFIVTRAGLYLYLYCQVHFDEGKAVYLKLD 180
```

```
Matches  241;   Conservative    0;   Mismatches    0;   Indels    0;

Qy    44  RRGRGEPGTALLVPELALGLGLALACLGLLLAVVSLGSRASLSAQEPAQEELVAE
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     9  RRGRGEPGTALLVPELALGLGLALACLGLLLAVVSLGSRASLSAQEPAQEELVAE

Qy   104  SELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQDGAQ
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    69  SELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQDGAQ

Qy   164  TVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYLKLDLLV
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   129  TVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYLKLDLLV

Qy   224  LRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFLTYF
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   189  LRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFLTYF

Qy   284  H 284
            |
Db   249  H 249
```

```
RESULT 9
AAE00891
ID    AAE00891 standard; protein; 249 AA.
XX
AC    AAE00891;
XX
DT    04-JUL-2001  (first entry)
XX
DE    Human TREPA (TNF related endothelium proliferative agent).
XX
KW    Human; tumour necrosis factor; TNF; angiogenesis; wound healing;
KW    TNF related endothelium proliferative agent; tumour; metastasis;
KW    grafting; vulnerary.
XX
OS    Homo sapiens.
XX
FH    Key            Location/Qualifiers
FT    Domain         98..249
FT                   /label= Extracellular_domain
XX
PN    US6207642-B1.
XX
PD    27-MAR-2001.
XX
PF    26-JUN-1998;   98US-00105343.
XX
PR    12-FEB-1997;   97US-00798692.
PR    10-FEB-1998;   98US-00021706.
XX
PA    (ABBO ) ABBOTT LAB.
XX
PI    Wiley SR;
XX
DR    WPI; 2001-280760/29.
DR    N-PSDB; AAD04350.
XX
PT    Inducing angiogenesis in mammal at desired sites for promoting w
PT    healing, by administering soluble fragment of extracellular doma
PT    tumor necrosis factor related endothelium proliferative agent pr
XX
PS    Claim 1; Col 75-76; 53pp; English.
XX
CC    The present invention relates to extracellular signal molecules,
CC    particularly members of tumour necrosis factor (TNF) family mole
CC    designated as TREPA (TNF related endothelium proliferative agent
CC    soluble biologically active TREPA are used to treat TREPA-associ
CC    diseases, tumours or metastases. TREPA is used for inducing angi
CC    in human for promoting wound healing and for vascularising graft
CC    for successful grafting and to promote tissue grafts. The presen
CC    acid sequence is clone ID #690050 human TREPA
```

```
EDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQD 155
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
EDQDPSELNPQTEESQDPAPFLNRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGQD 120

AGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYLKLD 215
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYLKLD 180

DGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFL 275
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFL 240

GLFQVH 284
||||||
GLFQVH 249
```

```
ndard; protein; 249 AA.

(first entry)

endothelium proliferative agent protein.

lium proliferative agent; TREPA; wound healing; cancer;
ing; vascularisation; apoptosis; autoimmune; birth control.

rT LAB.

98WO-US002859.

97US-00798692.
98US-00021706.

7255/38.
7613.

cleic acid encoding TREPA - useful for diagnosis and
autoimmune disease, tumours and inflammation.

ge 123-4; 142pp; English.

ted endothelium proliferative agent (TREPA), or its
r agonists, are used to treat a deficit of TREPA, e.g. to
d healing or tissue grafting, by promoting vascularisation,
ce apoptosis for treating cancer and eliminating autoreactive
an adjunct to cancer chemotherapy or antiviral treatment.
es can also be used to target cytotoxic agents or for
lation of the corresponding receptor, the nucleic acid for
used to transform tumour cells to render them more
o TREPA and to screen for TREPA mimics. Ribozymes, antisense
dies or peptides, are used to treat TREPA-associated
g. tumours and metastases (by inhibiting vascularisation),
or a wide range of autoimmune conditions, conditions
normal stimulation of epithelial cells (e.g.
sis), for birth control (inhibiting ovulation and placental
r other angiogenic conditions (e.g. ulcers)

AA;

      84.9%;   Score 241;   DB 2;   Length 249;
larity  100.0%;   Pred. No. 5.9e-213;
```

.A;

rity 100.0%; Score 241; DB 4; Length 249;
84.9%; Pred. No. 5.9e-213;
nservative 0; Mismatches 0; Indels 0; Gaps 0;

```
GEPGTALLVPLALGLGLALACLGLLLAVVSLGSRASLSAQEPAQEELVAEEDQDP 103
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
GEPGTALLVPLALGLGLALACLGLLLAVVSLGSRASLSAQEPAQEELVAEEDQDP 68

PQTEESQDPAPFLNRLVRPPRRSAPKGRKTRARRAIAAHYEVHPRPGQDGAQAGVDG 163
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
PQTEESQDPAPFLNRLVRPPRRSAPKGRKTRARRAIAAHYEVHPRPGQDGAQAGVDG 128

EEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYLKLDLLVDGVLA 223
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
EEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQVHFDEGKAVYLKLDLLVDGVLA 188

EFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFLTYFGLFQV 283
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
EFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFLTYFGLFQV 248
```

ard; protein; 273 AA.

first entry)

lular domain-containing fusion protein.

lular domain; tumour necrosis factor; TNF; angiogenesis;
ularisation; diabetic retinopathy; neovascular glaucoma;
; retinopathy of prematurity; retrolental fibroplasia;
tis; macular degeneration; arthritis; rheumatism;
neovascularisation: psoriasis; metastatic condition;
ur; sarcoma; carcinoma; benign tumour; haemophilic joint;
condition; myocardial angiogenesis; wound granulation;
ascular adhesion; telangiectasia; ischaemia: human;
c plaque neovascularisation; coronary atherosclerosis;
erosclerosis; pDC409-LZ-TWEAK; TWEAK receptor; TWEAKR;

000WO-US034755.

99US-0172878P.
000US-0203347P.

X CORP.

75/44.

64.

iogenesis in a mammal for treating diseases mediated by
e.g. solid tumors and vascular deficiencies of cardiac or
sue, by administering antagonist or agonist of TWEAK

XX
PS Example 1; Page 41; 46pp; English.
XX
CC The sequence represents a a fusion protein encoded by the express
CC vector pDC409-LZ-TWEAK. The fusion protein comprises a growth hor
CC leader, a leucine zipper multimerisation domain, and the extracel
CC domain of human TWEAK. The fusion protein was used in the isolati
CC human TWEAK receptor (TWEAKR)-expressing clones from a COS cell h
CC cDNA library. The TWEAK protein is a member of the tumour necrosi
CC (TNF) family and induces angiogenesis. TWEAKR may therefore be us
CC screen for and develop TWEAK agonists and antagonists for the mo
CC of angiogenesis, to be used in the treatment and diagnosis of hum
CC disease. The disorders mediated by angiogenesis include ocular di
CC characterised by ocular neovascularisation such as diabetic retin
CC neovascular glaucoma, retinopathy, retinoblastoma, retinopathy of prematurity,
CC retrolental fibroplasia, rubeosis, uveitis, macular degeneration
CC corneal graft neovascularisation, and inflammatory diseases such
CC arthritis, rheumatism and psoriasis. Other treatable diseases inc
CC malignant and metastatic conditions such as sarcomas and carcinom
CC benign tumours and preneoplastic conditions, myocardial angiogene
CC haemophilic joints, scleroderma, vascular adhesions, atherosclero
CC plaque neovascularisation, telangiectasia, wound granulation, cor
CC atherosclerosis, peripheral atherosclerosis and ischaemia
XX
SQ Sequence 273 AA;

Query Match 72.9%; Score 207; DB 4; Length 273;
Best Local Similarity 100.0%; Pred. No. 1.2e-181;
Matches 207; Conservative 0; Mismatches 0; Indels 0; G

```
Qy   78 SLGSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQDPAPFLNRLVRPPRRSAPKG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67 SLGSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQDPAPFLNRLVRPPRRSAPKG

Qy  138 ARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRA
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  127 ARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRA

Qy  198 LYCQVHFDEGKAVYLKLDLLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALI
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  187 LYCQVHFDEGKAVYLKLDLLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALI

Qy  258 SLRIRTLPWAHLKAAPFLTYFGLFQVH 284
          ||||||||||||||||||||||||||
Db  247 SLRIRTLPWAHLKAAPFLTYFGLFQVH 273
```

RESULT 11
AAE00895
ID AAE00895 standard; protein; 146 AA.
XX
AC AAE00895;
XX
DT 04-JUL-2001 (first entry)
XX
DE Human TREPA (TNF related endothelium proliferative agent) fragment
XX
KW Human; tumour necrosis factor; TNF; angiogenesis; wound healing;
KW TNF related endothelium proliferative agent; tumour; metastasis;
KW grafting; vulnerary.
XX
OS Homo sapiens.
XX
PN US6207642-B1.
XX
PD 27-MAR-2001.
XX
PF 26-JUN-1998; 98US-00105343.
XX
PR 12-FEB-1997; 97US-00798692.
PR 10-FEB-1998; 98US-00021706.
XX
PA (ABBO) ABBOTT LAB.

XX
CC The TNF-related endothelium proliferative agent (TREPA), or its
CC activators or agonists, are used to treat a deficit of TREPA, e.
CC promote wound healing or tissue grafting, by promoting vasculari
CC also to induce apoptosis for treating cancer and eliminating aut
CC T cells, as an adjunct to cancer chemotherapy or antiviral treat
CC TREPA peptides can also be used to target cytotoxic agents or fo
CC affinity isolation of the corresponding receptor, the nucleic ac
CC which can be used to transform tumour cells to render them more
CC responsive to TREPA and to screen for TREPA mimics. Ribozymes, a
CC RNA , antibodies or peptides, are used to treat TREPA-associated
CC diseases, e.g. tumours and metastases (by inhibiting vascularisa
CC inflammation or a wide range of autoimmune conditions, condition
CC involving abnormal stimulation of epithelial cells (e.g.
CC atherosclerosis), for birth control (inhibiting ovulation and pl
CC formation) or other angiogenic conditions (e.g. ulcers)
XX
SQ Sequence 189 AA;

Query Match 50.4%; Score 143; DB 2; Length 189;
Best Local Similarity 100.0%; Pred. No. 5.7e-123;
Matches 143; Conservative 0; Mismatches 0; Indels 0;

```
Qy  142 IAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLY
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   47 IAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLY

Qy  202 VHFDEGKAVYLKLDLLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPG
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  107 VHFDEGKAVYLKLDLLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPG

Qy  262 RTLPWAHLKAAPFLTYFGLFQVH 284
        |||||||||||||||||||||||
Db  167 RTLPWAHLKAAPFLTYFGLFQVH 189
```

RESULT 13
AAE00892
ID AAE00892 standard; protein; 189 AA.
XX
AC AAE00892;
XX
DT 04-JUL-2001 (first entry)
XX
DE Human UL4flag TREPA soluble construct.
XX
KW Human; tumour necrosis factor; TNF; angiogenesis; wound healing;
KW TREPA; TNF related endothelium proliferative agent; metastasis;
KW vulnerary; HUVEC; human umbilical vein endothelial cell; UL4flag
XX
OS Homo sapiens.
XX
PN US6207642-B1.
XX
PD 27-MAR-2001.
XX
PF 26-JUN-1998; 98US-00105343.
XX
PR 12-FEB-1997; 97US-00798692.
PR 10-FEB-1998; 98US-00021706.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Wiley SR;
XX
DR WPI; 2001-280760/29.
XX
PT Inducing angiogenesis in mammal at desired sites for promoting w
PT healing, by administering soluble fragment of extracellular doma
PT tumor necrosis factor related endothelium proliferative agent pr
XX
PS Example 2; Col 75-78; 53pp; English.
XX

0760/29.

iogenesis in mammal at desired sites for promoting wound
administering soluble fragment of extracellular domain of
is factor related endothelium proliferative agent protein.

Fig 1; 53pp; English.

invention relates to extracellular signal molecules,
members of tumour necrosis factor ('TNF') family molecules
a TREPA (TNF related endothelium proliferative agent).
ogically active TREPA are used to treat TREPA-associated
nours or metastases. TREPA is used for inducing angiogenesis
promoting wound healing and for vascularising grafted tissue
il grafting and to promote tissue grafts. The present amino
a is human TREPA fragment

AA;

larity 51.4%; Score 146; DB 4; Length 146;
  100.0%; Pred. No. 7.8e-126;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
IAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYL 198
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
IAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYL  60

VHFDEGKAVYLKLDLLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSS 258
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
VHFDEGKAVYLKLDLLVDGVLALRCLEEFSATAASSLGPQLRLCQVSGLLALRPGSS 120

RTLPWAHLKAAPFLTYFGLFQVH 284
|||||||||||||||||||||||
RTLPWAHLKAAPFLTYFGLFQVH 146
```

dard; protein; 189 AA.

(first entry)

endothelium proliferative agent protein 2.

lium proliferative agent; TREPA; wound healing; cancer;
ing; vascularisation; apoptosis; autoimmune; birth control.

98WO-US002859.

97US-00798692.
98US-00021706.

PT LAB.

7255/38.

cleic acid encoding TREPA - useful for diagnosis and
autoimmune disease, tumours and inflammation.

ge 125-6; 142pp; English.

nvention relates to extracellular signal molecules,
members of tumour necrosis factor (TNF) family molecules
TREPA (TNF related endothelium proliferative agent).
gically active TREPA are used to treat TREPA-associated
ours or metastases. TREPA is used for inducing angiogenesis
promoting wound healing and for vascularising grafted tissue
l grafting and to promote tissue grafts. The present amino
is human UU4flag TREPA soluble construct. This sequence
ologically active molecule is capable of inducing
in HUVEC (human umbilical vein endothelial cells) cells

AA;

arity   50.4%;   Score 143; DB 4; Length 189;
arity 100.0%;   Pred. No. 5.7e-123;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

YEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQ 201
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
YEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTRAGLYYLYCQ 106

EGKAVYLKLDLLVDGVLALRCLEBFSATAASSLGPQLRLCQVSGLLALRPGSSLRI 261
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
EGKAVYLKLDLLVDGVLALRCLEBFSATAASSLGPQLRLCQVSGLLALRPGSSLRI 166

WAHLKAAPFLTYFGLFQVH 284
|||||||||||||||||||
WAHLKAAPFLTYFGLFQVH 189

dard; protein; 208 AA.

(first entry)

rotein.

is factor receptor; signal transducer molecule; TNF; APO4;
abnormality; gestational abnormality; prostate cancer;
PO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
omain; immunogen; antibody preparation; breast carcinoma;
nan.

13A; 156pp; English.

98WO-US018393.

97US-00924634.

WASHINGTON.

191/17.
124.

rosis Factor family receptor polypeptides and ligands -
agnosis and treatment of prostate cancer and developmental
l abnormalities.

n describes isolated Tumor Necrosis Factor (TNF) family
peptides: APO4, APO6, APO8 and APO9 or their active
d isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
fragments. APO4 is useful for diagnosing prostate cancer by

CC   determining levels of APO4 in an individual. Prostate cancer can
CC   treated using APO4 selective binding agents linked to a therapeut
CC   moiety. APO4 polypeptides are also useful for identifying select
CC   binding agents, useful in diagnosis/treatment of disease by bind
CC   agents to the polypeptide/active fragment which is extracellular,
CC   expressed on the cell surface. The binding is preferably performe
CC   vivo. APO4 polypeptides/ active fragments are also useful for sc
CC   for agonists and antagonists by binding and observing the change
CC   activity. Effective pharmacological agents useful in diagnosis o
CC   treatment of disease are also identified using APO4 polypeptides/
CC   fragments and APO4 signal transducer molecules that specifically
CC   with a cytoplasmic domain of APO4 and detecting a change in level
CC   activity. The method is performed in vivo or in vitro. APO polype
CC   are all useful as immunogens for preparing antibodies. APO4 is al
CC   useful for diagnosis/treatment of developmental or gestational
CC   abnormalities. APO8 was transfected to human breast carcinoma cel
CC   MCF-7, and induced apoptosis
XX
SQ   Sequence 208 AA;

Query Match   37.7%; Score 107; DB 2; Length 208;
Best Local Similarity   99.5%; Pred. No. 7.8e-90;
Matches 207; Conservative   0; Mismatches   1; Indels   0; C

Qy   77 VSLGSRASLSAQEPAQEELVAEEDQDPSELNPQTEBSQDPAPPLNRLVRPRRSAP
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 VSLGSRASLSAQEPAQEELVAEEDQDPSELNPQTEBSQDPAPPLNRLVRPRRSAP

Qy   137 RARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFIVTF
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 RARRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEEARINSSSPLRYYRQIGEFIVTF

Qy   197 YLYCQVHFDEGKAVYLKLDLLVDGVLALRCLEBFSATAASSLGPQLRLCQVSGLLA
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 YLYCQVHFDEGKAVYLKLDLLVDGVLALRCLEBFSATAASSLGPQLRLCQVSGLLA

Qy   257 SSLRIRTLPWAHLKAAPFLTYFGLFQVH 284
         |||||||||||||||||||||||||||
Db   181 SSLRIRTLPWAHLKAAPFLTYFGLFQVH 208

RESULT 15
AAW93591
ID   AAW93591 standard; protein; 211 AA.
XX
AC   AAW93591;
XX
DT   18-JUN-1999 (first entry)
XX
DE   Mouse TNRL3 protein.
XX
KW   Tumour necrosis factor receptor; signal transducer molecule; TNF;
KW   developmental abnormality; gestational abnormality; prostate ca
KW   APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy;
KW   cytoplasmic domain; immunogen; antibody preparation; breast carci
KW   apoptosis; mouse.
XX
OS   Mus sp.
XX
PN   WO9911791-A2.
XX
PD   11-MAR-1999.
XX
PF   04-SEP-1998;   98WO-US018393.
XX
PR   05-SEP-1997;   97US-00924634.
XX
PA   (UNIW) UNIV WASHINGTON.
XX
PI   Chaudhary PM;
XX
DR   WPI; 1999-205191/17.
DR   N-PSDB; AAX23425.

crosis Factor family receptor polypeptides and ligands -
iagnosis and treatment of prostate cancer and developmental
al abnormalities.

g 13B; 156pp; English.

on describes isolated Tumor Necrosis Factor (TNF) family
ypeptides: APO4, APO6, APO8 and APO9 or their active
nd isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
fragments. APO4 is useful for diagnosing prostate cancer by
levels of APO4 in an individual. Prostate cancer can also be
g APO4 selective binding agents linked to a therapeutic
polypeptides are also useful for identifying selective
ts, useful in diagnosis/treatment of disease by binding of
e polypeptide/active fragment which is extracellular, or
the cell surface. The binding is preferably performed in
olypeptides/ active fragments are also useful for screening
and antagonists by binding and observing the changer in APO4
fective pharmacological agents useful in diagnosis or
disease are also identified using APO4 polypeptides/active
d APO4 signal transducer molecules that specifically interact
lasmic domain of APO4 and detecting a change in level of APO4
e method is performed in vivo or in vitro. APO polypeptides
ul as immunogens for preparing antibodies. APO4 is also
iagnosis/treatment of developmental or gestational
s. APO8 was transfected to human breast carcinoma cell line
nduced apoptosis

AA;

larity   16.2%;   Score 46;   DB 2;   Length 211;
Conservative   100.0%;   Pred.No. 9.6e-34;
Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

LRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 284
||||||||||||||||||||||||||||||||||||||||||||
LRLCQVSGLLALRPGSSLRIRTLPWAHLKAAPFLTYFGLFQVH 211

ndard; protein; 225 AA.

(first entry)

tumour necrosis factor related ligand (TRELL).

c necrosis factor related ligand; tnf; treatment; cancer;
isease; immune system; stimulation; suppression;
ion.

Location/Qualifiers
1. .21
/note= "hydrophobic, transmembrane domain"

97WO-US013945.

96US-0023541P.
96US-0028515P.
97US-0040820P.

EN INC.
GENEVA FACULTY MEDICINE.

PI	Chicheportiche Y, Browning JL;
XX	
DR	WPI; 1998-145619/13.
DR	N-PSDB; AAV18599.
XX	
PT	Tumour necrosis factor related ligand - useful for, e.g. treatir
PT	auto-immune disease and immune responses to tissue grafts.
PS	Claim 12; Page 48-50; 69pp; English.
XX	
CC	The sequence is that of mouse tumour necrosis factor related lig
CC	(TRELL). TRELL or active fragments can be included with a carrie
CC	pharmaceutical compositions to treat cancer, autoimmune diseases
CC	immune responses to tissue grafts, or to stimulate or suppress t
CC	system. It is useful to screen for TRELL receptors, by labelling
CC	detectable label and screening compositions for binding. Agents
CC	interfering with TRELL-receptor binding can also be screened for
CC	then be administered, optionally with interferon- gamma, to indu
CC	death or treat, suppress or alter immune responses (especially i
CC	human adenocarcinoma cells) involving a signal pathway between t
CC	its receptor. It's coding sequence can be used in gene therapy f
CC	related disorders in mammals (especially humans), e.g. tumours,
CC	autoimmune and inflammatory diseases or inherited genetic disord
CC	introducing into cells, and expressing, therapeutically effectiv
CC	of a vector, e.g. a virus comprising a gene encoding TRELL. It m
CC	be of use in the preparation of prepare probes for screening
CC	natural/synthetic DNAs for TRELL-encoding sequences and for anti
CC	therapy
XX	
SQ	Sequence 225 AA;

Query Match   11.3%;   Score 32;   DB 2;   Length 225;
Best Local Similarity   100.0%;   Pred. No. 7.6e-21;
Matches   32;   Conservative   0;   Mismatches   0;   Indels   0;

Qy   139 RRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEE 170
||||||||||||||||||||||||||||||||
Db   80 RRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEE 111

RESULT 17	
AAB07527	
ID	AAB07527 standard; protein; 225 AA.
XX	
AC	AAB07527;
XX	
DT	20-OCT-2000 (first entry)
XX	
DE	Amino acid sequence of a soluble recombinant murine TWEAK protei
XX	
KW	TWEAK protein; immunological disorder; immune response; inflamma
KW	TWEAK blocking agent; autoimmune disease; organ transplant rejec
KW	Graft-versus-Host disease; GVHD; lymphoid cell malignancy; shock
XX	
OS	Mus sp.
XX	
PN	WO20042073-A1.
XX	
PD	20-JUL-2000.
XX	
PF	14-JAN-2000; 2000WO-US001044.
XX	
PR	15-JAN-1999; 99US-0116168P.
XX	
PA	(BIOJ) BIOGEN INC.
XX	
PI	Rennert P;
XX	
DR	WPI; 2000-476036/41.
XX	
PT	Preventing and treating immune responses using modulators, espec
PT	antibodies, of TWEAK, TWEAK receptors and TWEAK ligands, useful :
PT	treating e.g. inflammation and graft versus host disease.

CC   skin disease; skeletal muscle disease; adipose tissue disease;
CC   gastrointestinal tract disease; pancreatic disease; reproductive
CC   disease; neural disease; cartilage disease; bone disease; connect
CC   tissue disease; cellular death; and a pathological condition of a
XX   expressing a TWEAK receptor.
SQ   Sequence 249 AA;

Query Match   11.3%; Score 32; DB 7; Length 249;
Best Local Similarity 100.0%; Pred. No. 8.4e-21;
Matches   32; Conservative   0; Mismatches   0; Indels   0; C

Qy   139 RRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEE 170
             |||||||||||||||||||||||||||||||||
Db   104 RRAIAAHYEVHPRPGQDGAQAGVDGTVSGWEE 135

RESULT 19
AAG01265
ID   AAG01265 standard; protein; 58 AA.
XX
AC   AAG01265;
XX
DT   06-OCT-2000 (first entry)
XX
DE   Human secreted protein, SEQ ID NO: 5346.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA iso
KW   gene therapy; chromosome mapping.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-00200610.
XX
PR   26-FEB-1999; 99US-0122487P.
XX
PA   (GEST) GENSET.
XX
PI   Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR   WPI; 2000-500387/45.
DR   N-PSDB; AAC01127..
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and fo
PT   diagnostic, forensic, gene therapy and chromosome mapping procedu
XX
PS   Claim 13; SEQ ID NO 5346; 71pp + Sequence Listing; English.
XX
CC   The present sequence is a polypeptide encoded by one of a large n
CC   5' ESTs derived from mRNAs encoding secreted proteins. The 5' EST
CC   prepared from total human RNAs or polyA+ RNAs derived from 30 dif
CC   tissues. EST sequences usually correspond mainly to the 3' untran
CC   region (UTR) of the mRNA because they are often obtained from oli
CC   primed cDNA libraries. Such ESTs are not well suited for isolatin
CC   sequences derived from the 5' ends of mRNAs and even in those cas
CC   longer cDNA sequences have been obtained, the full 5' UTR is rare
CC   included. 5' ESTs are derived from mRNAs with intact 5' ends and
CC   therefore be used to obtain full length cDNAs and genomic DNAs. 5
CC   are also used in diagnostic, forensic, gene therapy and chromosom
CC   mapping procedures. They are used to obtain upstream regulatory se
CC   and to design expression and secretion vectors
XX
SQ   Sequence 58 AA;

Query Match   3.2%; Score 9; DB 3; Length 58;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches   9; Conservative   0; Mismatches   0; Indels   0; G

XX
DE   Human DITHP receptor.
XX
KW   Human; dithp; diagnostic and therapeutic polynucleotide; diagnos
KW   cancer; cell proliferative disorder; autoimmune disorder;
KW   inflammatory disorder; infection; hormonal disorder; metabolic d
KW   neurological disorder; gastrointestinal disorder; transport diso
KW   connective tissue disorder; drug screening; proteome analysis;
KW   gene therapy; antisense therapy; genotyping; transgenic animal;
KW   disease model; toxicological testing; transcript imaging; recept
XX
OS   Homo sapiens.
XX
XX
PN   WO200297031-A2.
XX
PD   05-DEC-2002.
XX
PF   27-MAR-2002; 2002WO-US010056.
XX
PR   28-MAR-2001; 2001US-0279619P.
PR   29-MAR-2001; 2001US-0280067P.
PR   29-MAR-2001; 2001US-0280068P.
PR   16-MAY-2001; 2001US-0291280P.
PR   17-MAY-2001; 2001US-0291829P.
PR   17-MAY-2001; 2001US-0291849P.
PR   19-JUN-2001; 2001US-0299428P.
PR   20-JUN-2001; 2001US-0299776P.
PR   20-JUN-2001; 2001US-0300001P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
PI   Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;
PI   Daughtery, SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gersti
PI   Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris
PI   Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
XX
DR   WPI; 2003-123518/12.
DR   N-PSDB; ACC46177.
XX
PT   Novel human diagnostic and therapeutic polypeptide useful for id
PT   test compound which specifically binds to a polypeptide encoded
PT   diagnostic and therapeutic polynucleotide, and to induce antibod
XX
PS   Claim 27; SEQ ID NO 770; 591pp; English.
XX
CC   The invention relates to novel human diagnostic and therapeutic
CC   polynucleotides designated dithp (ACC46080-ACC46749) and to thei
CC   proteins (DITHP; ABR41136-ABR41812). The invention also relates
CC   polynucleotide sequences at least 90% identical to the dithp cDN
CC   sequences of the invention; recombinant vectors; host cells and
CC   transgenic organisms comprising a dithp nucleic acid sequence; t
CC   recombinant production of DITHP proteins; antibodies specific fo
CC   proteins; microarrays comprising dithp nucleic acid sequences; m
CC   detecting dithp nucleotide and protein sequences; methods of scr
CC   for compounds which specifically bind a DITHP protein; and metho
CC   assessing the toxicity of test compounds using a dithp hybridisa
CC   probe. Dithp nucleic acid sequences and DITHP proteins may be us
CC   diagnosis of a wide variety of conditions including cancer and o
CC   proliferative disorders; autoimmune or inflammatory disorders; b
CC   viral, fungal or parasitic infections; hormonal disorders; metab
CC   disorders; neurological disorders; gastrointestinal disorders; t
CC   disorders; and connective tissue disorders. They may also be use
CC   screen for modulators of protein activity or gene expression. DI
CC   proteins can additionally be used in analysis of the proteome of
CC   or cell type and to induce antibodies. The dithp nucleic acids a
CC   additionally useful in somatic or germline gene therapy of the d
CC   mentioned above, as a source of antisense sequences, as a source
CC   probes and primers, in genotyping and identification of individu
CC   the generation of transgenic animal models of human disease or k
CC   humanised animals, in toxicological testing, and in transcript i
CC   The present sequence represents a DITHP protein which has recept
CC   activity. Note: The sequence data for this patent did not form p

---

.DFEIS 9
|||||
.DFEIS 54

dard; protein; 58 AA.

(first entry)

ed protein, SEQ ID NO: 5347.

C; expressed sequence tag; secreted protein; cDNA isolation;
chromosome mapping.

2000EP-00200610.

99US-0122487P.

T.

dwards J, Duclert A, Giordano J;

381/45.
272.

cid that is a 5' expressed sequence tag (5' EST) for
DNAs and genomic DNAs that correspond to 5'ESTs and for
forensic, gene therapy and chromosome mapping procedures.

ID NO 5347; 71pp + Sequence Listing; English.

sequence is a polypeptide encoded by one of a large number of
ed from mRNAs encoding secreted proteins. The 5' ESTs were
total human RNAs or polyA+ RNAs derived from 30 different
sequences usually correspond mainly to the 3' untranslated
of the mRNA because they are often obtained from oligo-dT
libraries. Such ESTs are not well suited for isolating cDNA
ived from the 5' ends of mRNAs and even in those cases where
sequences have been obtained, the full 5' UTR is rarely
ESTs are derived from mRNAs with intact 5' ends and can
used to obtain full length cDNAs and genomic DNAs. 5' ESTs
l in diagnostic, forensic, gene therapy and chromosome
dures. They are used to obtain upstream regulatory sequences
expression and secretion vectors

A;

arity   3.2%;   Score 9;   DB 3;   Length 58;
  100.0%;   Pred.No. 2.9;
Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

.DFEIS 9
|||||
.DFEIS 54

dard; protein; 365 AA.

(first entry)

pecification, but was obtained in electronic format directly
ftp.wipo.int/pub/published_pct_sequences

AA;

```
arity        3.2%;   Score 9;   DB 6;   Length 365;
100.0%;   Pred. No. 17;
onservative   0;   Mismatches   0;   Indels   0;   Gaps   0;
```

```
GRRGE 50
|||||
GRRGE 30
```

dard; protein; 748 AA.

(first entry)

o acid sequence for GVs-9.

vaccae protein; antigen; T cell activation; cytokine;
l maturation; infectious disease; immune disorder; cancer;
ystem; mycobacterial infection; allergy; tuberculosis;
oidosis; lung cancer; asthma; skin disorder; psoriasis;
czema; alopecia areata; skin cancer; basal carcinoma;
carcinoma; melanoma.

vaccae.

98WO-NZ000189.

```
97US-00996624.
97US-00997080.
97US-00997362.
98US-00095855.
98US-00156181.
98US-00205426.
```

IS RES & DEV CORP LTD.

n J, Visser ES,   Skinner MA,   Prestidge RL;

163/36.
368.

une response to an antigen.

209-210; 243pp; English.

provides heat-killed Mycobacterium vaccae, or recombinant
teins. The M. vaccae proteins may be employed to activate T
ural killer cells, to stimulate the production of cytokines,
e expression of co-stimulatory molecules on dendritic cells
, and to enhance dendritic cell maturation and function. The
be expressed by standard recombinant methodology.
l compositions comprising the proteins or nucleic acid
oding the proteins can be used for the treatment,
nd detection of disorders including infectious diseases,
ers and cancer. In particular, the compounds and methods are
tment of diseases of the respiratory system, such as
infections, asthma, allergies, tuberculosis, leprosy,
nd lung cancers, and disorders of the skin such as
pic dermatitis, eczema, allergic contact dermatitis,
ta, and skin cancers such as basal carcinoma, squamous cell
melanoma

XX
SQ   Sequence 748 AA;

```
Query Match          3.2%;   Score 9;   DB 2;   Length 748;
Best Local Similarity   100.0%;   Pred. No. 34;
Matches        9;   Conservative   0;   Mismatches   0;   Indels   0;
```

```
Qy    60 ALGLGLALA 68
         |||||||||
Db   282 ALGLGLALA 290
```

RESULT 23
ABB73512
ID   ABB73512 standard; protein; 749 AA.
XX
AC   ABB73512;
XX
DT   08-APR-2002   (first entry)
XX
DE   M vaccae GVs-9 protein SEQ ID NO: 154.
XX
KW   Skin disorder; psoriasis; atopic dermatitis; allergic contact der
KW   alopecia areata; skin cancer; Mycobacterium vaccae; melanoma; cyt
KW   antipsoriatic; dermatological; antiinflammatory; antiallergic;
KW   Th2 immune response; immunomodulatory.
XX
OS   Mycobacterium vaccae.
XX
PN   US6328978-B1.
XX
PD   11-DEC-2001.
XX
PF   02-JUN-1999;   99US-00324542.
XX
PR   23-DEC-1997;   97US-00997080.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI   Watson JD,   Tan PLJ,   Prestidge R;
XX
DR   WPI; 2002-138361/18.
DR   N-PSDB; ABL36274.
XX
PT   Inhibiting skin inflammation associated with skin disorder e.g.
PT   psoriasis, by administering composition comprising delipidated ar
PT   deglycolipidated Mycobacterium vaccae cells or Mycobacterium vacc
PT   culture filtrate.
XX
PS   Example 6; Col 161-164; 116pp; English.
XX
CC   The present invention relates to a method of inhibiting skin infl
CC   associated with a skin disorder selected from psoriasis, atopic
CC   dermatitis and allergic contact dermatitis, which involves admini
CC   a composition containing delipidated and deglycolipidated Mycobac
CC   vaccae cells or M. vaccae culture filtrate. The skin disorder to
CC   treated may also include alopecia areata, and skin cancers such a
CC   cell carcinoma, squamous cell carcinoma and melanoma. The composi
CC   acts by inhibiting the Th2 immune response. The present sequence
CC   protein described in the exemplification of the invention
XX
SQ   Sequence 749 AA;

```
Query Match          3.2%;   Score 9;   DB 5;   Length 749;
Best Local Similarity   100.0%;   Pred. No. 34;
Matches        9;   Conservative   0;   Mismatches   0;   Indels   0;   G
```

```
Qy    60 ALGLGLALA 68
         |||||||||
Db   282 ALGLGLALA 290
```

RESULT 24

ABM48382
ID   ABM48382 standard; protein; 54 AA.
XX
AC   ABM48382;
XX
DT   20-OCT-2003 (first entry)
XX
DE   Propionibacterium acnes predicted ORF-encoded polypeptide #13058
XX
KW   Acne vulgaris; antiseborrhoeic; dermatological; antibacterial;
KW   immunostimulant; immune response; vaccine.
XX
OS   Propionibacterium acnes.
XX
PN   WO200303515-A1.
XX
PD   24-APR-2003.
XX
PF   11-OCT-2002; 2002WO-US032727.
XX
PR   15-OCT-2001; 2001US-00978825.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JI
PI   Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Cart
PI   Barth B, Vallieve-Douglass J;
XX
DR   WPI; 2003-381789/36.
DR   N-PSDB; ACF64481.
XX
PT   New Propionibacterium acnes polypeptides and polynucleotides enc
PT   polypeptide, useful for diagnosing, preventing or treating acne
PT   or for stimulating an immune response specific for a P. acnes pr
XX
PS   Example 1; SEQ ID NO 13058; 1481pp; English.
XX
CC   The invention relates to an isolated polynucleotide (ACF64435-AC
CC   encoding a Propionibacterium acnes protein. The invention also r
CC   polypeptides encoded by the polynucleotides (ABM35624-ABM64536)
CC   immunogenic fragments of P. acnes polypeptides. The invention
CC   additionally encompasses expression vectors and host cells compr
CC   polynucleotide of the invention; antibodies against polypeptides
CC   invention; fusion proteins comprising a polypeptide of the inven
CC   method for stimulating an immune response specific for a P. acne
CC   polypeptide and an isolated T cell population comprising T cells
CC   via this method, a vaccine composition (comprising P. acnes poly
CC   polynucleotides, antibodies, fusion proteins, T cell populations
CC   antigen-presenting cells that express the polypeptide); a method
CC   for detecting or determining the presence or absence of P. acnes
CC   patient; and a method for inhibiting the development of P. acnes
CC   patient. The P. acnes polypeptides, polynucleotides, antibodies,
CC   proteins, T cell populations or antigen-presenting cells that ex
CC   polypeptides are useful for diagnosing, preventing or treating a
CC   vulgaris, or for stimulating an immune response specific for a P
CC   protein. The polynucleotides can also be used as probes or prime
CC   nucleic acid hybridisation. The vaccine composition is useful fo
CC   stimulation of an immune response against P. acnes, or for treat
CC   and the kit is useful for performing a diagnostic assay. The pre
CC   sequence represents a polypeptide predicted to be encoded by an
CC   reading frame) contained within the P. acnes polynucleotides of
CC   invention. Note: The sequence data for this patent did not form
CC   the printed specification, but was obtained in electronic format
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 54 AA;

Query Match   2.8%; Score 8; DB 6; Length 54;
Best Local Similarity   100.0%; Pred. No. 23;
Matches   8; Conservative   0; Mismatches   0; Indels   0;

Qy   11 RRLPLPRS 18
               ||||||||

---

ndard; protein; 54 AA.

(first entry)

erium acnes immunogenic protein #12759.

me; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
ophthalmitis; bone; joint; central nervous system; ELISA;
lesion; acne vulgaris; enzyme linked immunosorbent assay;
al; osteopathic; neuroprotectant.

erium acnes.

A2.

2001WO-US012865.

2000US-0199047P.
2000US-0208841P.
2000US-0216747P.

XA CORP.

Persing DH, Mitcham JL, Wang SS, Bhatia A;
e J, Zhang Y, Jen S, Carter D;

6774/71.
9552.

erium acnes polypeptides and nucleic acids useful for
against and diagnosing infections, especially useful for
e vulgaris.

EQ ID NO 13058; 1069pp; English.

U39105-AAU68017 represent Propionibacterium acnes immunogenic
. The proteins and their associated DNA sequences are used in
t, prevention and diagnosis of medical conditions caused by
disorders include SAPHO syndrome (synovitis, acne,
hypertosis and osteomyelitis), uveitis and endophthalmitis.
also involved in infections of bone, joints and the central
em, however it is particularly involved in the inflammatory
ciated with acne vulgaris. A method for detecting the
absence of P. acnes in a patient comprises contacting a
a binding agent that binds to the proteins of the invention
ing the amount of bound protein in the sample. The
may be used as antigens in the production of antibodies
P. acnes proteins. These antibodies can be used to
expression and activity of P. acnes polypeptides and
eat P. acnes infections. The antibodies may also be used as
gents for determining P. acnes presence, for example, by
d immunosorbent assay (ELISA). Note: The sequence data for
did not form part of the printed specification, but was
electronic format directly from WIPO at
/pub/published_pct_sequences

AA;

larity 100.0%; Pred. No. 23;
Conservative   0; Mismatches   0; Indels   0; Gaps   0;

PLPRS 18
||||
PLPRS 28

Db

LPRS 28

dard; protein; 55 AA.
XX
(first entry)
XX
DE   encoded by probe for measuring cervical gene expression.
XX
KW   microarray; gene expression; cervical epithelial cell;
KW   er.
XX
2.
XX
2001WO-US000670.
XX
2000US-0180312P.
2000US-0207456P.
2000US-0608408.
2000US-0632366.
2000US-0234687P.
2000US-0236359P.
2000GB-00024263.
XX
ULAR DYNAMICS INC.
XX
zel DK,   Chen W,   Rank DR;
XX
901/53.
XX
derived single exon nucleic acid probes useful for analyzing
on in human cervical epithelial cells.
XX
ID NO 25684; 487pp; English.
XX
nvention relates to human single exon nucleic acid probes
I10068-AAI28459). The present sequence is a peptide encoded
robe. The SENPs are derived from human HeLa cells. The SENPs
o produce a single exon microarray, which can be used for
an gene expression in a sample derived from human cervical
lls. By measuring gene expression, the probes are therefore
ding and/or staging of diseases of the cervix, notably
er. Note:The sequence data for this patent did not form
rinted specification, but was obtained in electronic format
WIPO at ftp.wipo.int/pub/published_pct_sequences
a;

arity   2.8%;   Score 8;   DB 4;   Length 55;
100.0%;   Pred. No. 23;
onservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

LALA 68
|||||
LALA 19

dard; peptide; 55 AA.
XX
(first entry)

DE   Peptide #10143 encoded by human foetal liver single exon probe.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid p:
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,   Hanzel DK,   Chen W,   Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for a
PT   gene expression in human fetal liver.
XX
PS   Claim 27; SEQ ID NO 35272; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for mea
CC   human gene expression in a sample derived from human foetal liver
CC   single exon nucleic acid probes may be used for predicting, meas
CC   displaying gene expression in samples derived from human fetal li
CC   present sequence is a peptide encoded by a single exon nucleic ac
CC   of the invention. Note:The sequence data for this patent did not
CC   part of the printed specification, but was obtained in electronic
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 55 AA;

Query Match   2.8%;   Score 8;   DB 4;   Length 55;
Best Local Similarity   100.0%;   Pred. No. 23;
Matches   8;   Conservative   0;   Mismatches   0;   Indels   0;   G

Qy   61 LGLGLALA 68
            ||||||||
Db   12 LGLGLALA 19

RESULT 28
AAM36451
ID   AAM36451 standard; protein; 55 AA.
XX
AC   AAM36451;
XX
DT   17-OCT-2001 (first entry)
XX
DE   Peptide #10488 encoded by probe for measuring placental gene expr
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.

PT hearts.
XX
PS Claim 15; SEQ ID NO 27753; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probe
CC measuring human gene expression in a sample derived from human
CC ABA21535-ABA41305). The present sequence is a protein encoded by
CC probe. The probes may be used for predicting, measuring and disp
CC gene expression in samples derived from the human heart via mic
CC By measuring gene expression, the probes are useful for predict
CC diagnosing, grading, staging, monitoring and prognosing disease
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly f
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 55 AA;

Query Match   2.8%; Score 8; DB 4; Length 55;
Best Local Similarity 100.0%; Pred. No. 23;
Matches   8; Conservative   0; Mismatches   0; Indels   0;

Qy   61 LGLGLALA 68
         ||||||||
Db   12 LGLGLALA 19

RESULT 30
AAM76342
ID AAM76342 standard; protein; 55 AA.
XX
AC AAM76342;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 36
XX
KW Human; bone marrow expressed exon; gene expression analysis; pr
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 36648; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic a
CC probes which are derived from genomic sequences expressed in the
CC bone marrow. They can be used to measure gene expression in bone
CC samples, which may enable the improved diagnosis and treatment o
CC such as lymphoma, leukaemia and myeloma. The present sequence is

---

2000US-0207456P.
2000US-00608408.
2000US-00632366.
2000US-0234687P.
2000US-0236359P.
2000GB-00024263.

CULAR DYNAMICS INC.

nzel DK, Chen W, Rank DR;

8897/53.

-derived single exon nucleic acid probes useful for analyzing
ion in human placenta.

Q ID NO 36720; 654pp; English.

invention relates to single exon nucleic acid probes (SENP:
-AAI57546). The present sequence is a peptide encoded by one
The probes are useful for producing a microarray for
measuring and displaying gene expression in samples derived
placenta. The probes are useful for antenatal diagnosis of
c disorders

AA;

larity 100.0%; Pred. No. 23;
Conservative   0; Mismatches   0; Indels   0; Gaps   0;

GLALA 68
|||||
GLALA 19

ndard; protein; 55 AA.

(first entry)

2 encoded by probe for measuring heart cell gene expression.

expression; heart; microarray; vascular system;
ar disease; hypertension; cardiac arrhythmia;
eart disease.

A2.

(first entry)

2001WO-US000666.

2000US-0180312P.
2000US-0207456P.
2000US-00608408.
2000US-00632366.
2000US-0234687P.
2000US-0236359P.
2000GB-00024263.

CULAR DYNAMICS INC.

nzel DK, Chen W, Rank DR;

8899/53.

nucleic acid probes for analyzing gene expression in human

Left column (fragment, cut off at left edge)

```
d by one of the probes of the invention

i;
       2.8%;  Score 8;  DB 4;  Length 55;
rity  100.0%;  Pred. No. 23;
onservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

ALA 68
|||
ALA 19

ard; protein; 55 AA.

(first entry)

xpressed single exon probe encoded protein SEQ ID NO: 35633.

expressed exon; gene expression analysis; probe; microarray;
isease; multiple sclerosis; schizophrenia; epilepsy; cancer.

2.

2001WO-US000667.

2000US-0180312P.
2000US-0207456P.
2000US-00608408.
2000US-0632366.
2000US-0234687P.
2000US-0236359P.
2000GB-0024263.

ULAR DYNAMICS INC.

zel DK, Chen W, Rank DR;

446/52.

ucleic acid probes for analyzing gene expression in human

Q ID NO 35633; 650pp + Sequence Listing; English.

nvention provides a number of single exon nucleic acid
are derived from genomic sequences expressed in the human
an be used to measure gene expression in brain cell samples,
ble the diagnosis and improved treatment of nervous system
as Alzheimer's disease, multiple sclerosis, schizophrenia,
cancers. The present sequence is a protein encoded by one of
the invention

A;

       2.8%;  Score 8;  DB 4;  Length 55;
arity  100.0%;  Pred. No. 23;
onservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

LALA 68
|||
LALA 19
```

Right column

```
ABG58050
ID   ABG58050 standard; peptide; 55 AA.
XX
AC   ABG58050;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver peptide, SEQ ID No 36698.
XX
KW   Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
KW   hypercholesterolaemia; coronary heart disease.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488898/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for a
PT   gene expression in human adult liver.
XX
PS   Claim 27; SEQ ID NO 36698; 658pp; English.
XX
CC   The invention relates to a single exon nucleic acid probe (SENP)
CC   measuring human gene expression in a sample derived from human ad
CC   liver, comprising one of 13109 defined nucleotide sequences give
CC   specification (or complements/ fragments).  The probe hybridises a
CC   stringency to a nucleic acid molecule expressed in the human adul
CC   (I) may be used for predicting, measuring and displaying gene exp
CC   in samples derived from human adult liver.  The genes identified t
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia
CC   associated with coronary heart disease.  ABG47348-ABG59930 repres
CC   liver single exon encoded peptides of the invention.  Note: The se
CC   information for this patent does not appear in the printed specif
CC   but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 55 AA;

Query Match          2.8%;  Score 8;  DB 4;  Length 55;
Best Local Similarity  100.0%;  Pred. No. 23;
Matches    8;  Conservative   0;  Mismatches   0;  Indels   0;  C

Qy      61 LGLGLALA 68
            ||||||||
Db      12 LGLGLALA 19

RESULT 33
ABG45635
ID   ABG45635 standard; peptide; 55 AA.
XX
AC   ABG45635;
XX
DT   19-AUG-2002  (first entry)
XX
```

e encoded by genome-derived single exon probe SEQ ID 35300.

e exon probe; asthma; lung cancer; COPD; ILD;
ructive pulmonary disease; interstitial lung disease;
opathic pulmonary fibrosis; neurofibromatosis;
erosis; Gaucher's disease; Niemann-Pick disease;
dlak syndrome; sarcoidosis; pulmonary haemosiderosis;
stiocytosis; lymphangioleiomyomtosis; Kartagener syndrome;
veolar proteinosis; fibrocystic pulmonary dysplasia;
ary dyskinesis; pulmonary hypertension;
rane disease.

A2.

2001WO-US000665.

2000US-0180312P.
2000US-0207456P.
2000US-00608408.
2000US-00632366.
2000US-0234687P.
2000US-0236359P.
2000GB-00024263.

CULAR DYNAMICS INC.

zel DK,   Chen W,   Rank DR;

183/15.

ressable set of single exon nucleic acid probes, used to
expression in human lung samples.

Q ID NO 35300; 634pp; English.

n relates to a spatially-addressable set of single exon
probes for measuring gene expression in a sample derived
ing comprising single exon nucleic acid probes having one of
c acid sequences mentioned in the specification, or their
or the 12387 open reading frames derived from the 12614
included are a microarray comprising the novel set of probes
ad in the human lung; measuring gene expression to a nucleic
ad in the human lung; measuring gene expression in a sample
human lung, comprising (a) contacting the array with a
detectably labeled nucleic acids derived from human lung
measuring the label detectably bound to each probe of the
fying exons in a eukaryotic genome, comprising (a)
ly predicting at least one exon from genomic sequences of
; and (b) detecting specific hybridisation of detectably
ic acids from eukaryote lung mRNA, to a single exon probe,
ment identical to the predicted exon, the probe is included
mentioned microarray; assigning exons to a single gene,
) identifying exons from genomic sequence by the method
measuring the expression of each of the exons in several
r cell types using hybridisation to a single exon
having a probe with the exon, where a common pattern of
the exons in the tissues and/or cell types indicates that
uld be assigned to a single gene; a peptide comprising one
ences, mentioned in the specification, or encoded by the
reading frames (ORF). The Probes are used for gene expression
l for identifying exons in a gene, particularly using human
mRNA and for the study of lung diseases such as asthma, lung
nic obstructive pulmonary disease (COPD), interstitial lung
, familial idiopathic pulmonary fibrosis, neurofibromatosis,
erosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
me, sarcoidosis, pulmonary haemosiderosis, pulmonary
i, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
drome, fibrocystic pulmonary dysplasia, primary ciliary
pulmonary hypertension and hyaline membrane disease. The

CC present sequence is a peptide/protein encoded by a single exon
CC the invention. Note: The sequence data for this patent did not f
CC of the printed specification, but was obtained in electronic for
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 55 AA;

Query Match   2.8%; Score 8; DB 5; Length 55;
Best Local Similarity 100.0%; Pred. No. 23;
Matches   8; Conservative 0; Mismatches 0; Indels 0;

Qy   61 LGLGLALA 68
         ||||||||
Db   12 LGLGLALA 19

RESULT 34
AAM21621
ID   AAM21621 standard; protein; 65 AA.
XX
AC   AAM21621;
XX
DT   12-OCT-2001 (first entry)
XX
DE   Peptide #8055 encoded by probe for measuring cervical gene expre
XX
KW   Probe; human; microarray; gene expression; cervical epithelial c
KW   cervical cancer.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,   Hanzel DK,   Chen W,   Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 27; SEQ ID NO 26447; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid
CC   (SENP: see AAI10068-AAI28459). The present sequence is a peptide
CC   by one such probe. The SENPs are derived from human HeLa cells.
CC   can be used to produce a single exon microarray, which can be us
CC   measuring human gene expression in a sample derived from human c
CC   epithelial cells. By measuring gene expression, the probes are t
CC   useful in grading and/or staging of diseases of the cervix, nota
CC   cervical cancer. Note: The sequence data for this patent did not
CC   part of the printed specification, but was obtained in electronic
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 65 AA;

Query Match   2.8%; Score 8; DB 4; Length 65;
Best Local Similarity 100.0%; Pred. No. 27;
Matches   8; Conservative 0; Mismatches 0; Indels 0;

```
LALA 68
||||
LALA 29

dard; peptide; 65 AA.

(first entry)

7 encoded by human foetal liver single exon probe.

liver; gene expression; single exon nucleic acid probe.

2.

2001WO-US000669.

2000US-0180312P.
2000US-0207456P.
2000US-00608408.
2000US-00612366.
2000US-0234687P.
2000US-0236359P.
2000GB-00024263.

JLAR DYNAMICS INC.

zel DK,   Chen W,   Rank DR;

147/52.

derived single exon nucleic acid probes useful for analyzing
on in human fetal liver.

ID NO 36616; 639pp + Sequence Listing; English.

relates to a single exon nucleic acid probe for measuring
pression in a sample derived from human foetal liver. The
cleic acid probes may be used for predicting, measuring and
e expression in samples derived from human fetal liver. The
ce is a peptide encoded by a single exon nucleic acid probe
on. Note: The sequence data for this patent did not form
inted specification, but was obtained in electronic format
WIPO at ftp.wipo.int/pub/published_pct_sequences

;

urity       2.8%;   Score 8;   DB 4;   Length 65;
100.0%;   Pred. No. 27;
nservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

ALA 68
|||
ALA 29

lard; protein; 65 AA.

(first entry)

) encoded by probe for measuring placental gene expression.
```

```
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000;  2000US-0180312P.
PR   26-MAY-2000;  2000US-0207456P.
PR   30-JUN-2000;  2000US-00608408.
PR   03-AUG-2000;  2000US-00632366.
PR   21-SEP-2000;  2000US-0234687P.
PR   27-SEP-2000;  2000US-0236359P.
PR   04-OCT-2000;  2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,   Chen W,   Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for a
PT   gene expression in human placenta.
XX
PS   Claim 27; SEQ ID NO 38192; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes
CC   see AAI3115-AAI57546). The present sequence is a peptide encoded
CC   such probe. The probes are useful for producing a microarray for
CC   predicting, measuring and displaying gene expression in samples d
CC   from human placenta. The probes are useful for antenatal diagnosi
CC   human genetic disorders
XX
SQ   Sequence 65 AA;

Query Match           2.8%;     Score 8;    DB 4;    Length 65;
Best Local Similarity   100.0%;    Pred. No. 27;
Matches      8;   Conservative   0;   Mismatches   0;   Indels   0;   G

Qy     61 LGLGLALA 68
           ||||||||
Db     22 LGLGLALA 29

RESULT 37
ABB26890
ID   ABB26890 standard; protein; 65 AA.
XX
AC   ABB26890;
XX
DT   23-JAN-2002 (first entry)
XX
DE   Protein #8889 encoded by probe for measuring heart cell gene expr
XX
KW   Human; gene expression; heart; microarray; vascular system;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000666.
XX
PR   04-FEB-2000;  2000US-0180312P.
PR   26-MAY-2000;  20COUS-0207456P.
```

XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   gene expression in human bone marrow.
XX
PS   Example 4; SEQ ID NO 38012; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic a
CC   probes which are derived from genomic sequences expressed in the
CC   bone marrow. They can be used to measure gene expression in bone
CC   samples, which may enable the improved diagnosis and treatment o
CC   such as lymphoma, leukaemia and myeloma. The present sequence is
CC   protein encoded by one of the probes of the invention
XX
SQ   Sequence 65 AA.

Query Match   2.8%; Score 8; DB 4;   Length 65;
Best Local Similarity   100.0%; Pred. No. 27;
Matches   8; Conservative   0; Mismatches   0; Indels   0;

Qy   61 LGLGLALA 68
            ||||||||
Db   22 LGLGLALA 29

RESULT 39
AAM64984
ID   AAM64984 standard; protein; 65 AA.
XX
AC   AAM64984;
XX
DT   05-NOV-2001 (first entry)
XX
DE   Human brain expressed single exon probe encoded protein SEQ ID N
XX
KW   Human; brain expressed exon; gene expression analysis; probe; mi
KW   Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000667.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in
PT   brains.
XX
PS   Example 4; SEQ ID NO 37089; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic ac
CC   probes which are derived from genomic sequences expressed in the
CC   brain. They can be used to measure gene expression in brain cell
CC   which may enable the diagnosis and improved treatment of nervous
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizop
CC   epilepsy and cancers. The present sequence is a protein encoded b
CC   the probes of the invention

2000US-00608408.
2000US-00632366.
2000US-0234687P.
2000US-0236359P.
2000GB-00024263.

CULAR DYNAMICS INC.

zel DK, Chen W, Rank DR;

8899/53.

nucleic acid probes for analyzing gene expression in human

Q ID NO 28660; 530pp; English.

invention relates to single exon nucleic acid probes for
nan gene expression in a sample derived from human heart (see
41305). The present sequence is a protein encoded by one such
obes may be used for predicting, measuring and displaying
ion in samples derived from the human heart via microarrays.
gene expression, the probes are useful for predicting,
rading, staging, monitoring and prognosing diseases of the
and vascular system e.g. cardiovascular disease,
. cardiac arrhythmias and congenital heart disease. Note: The
for this patent did not form part of the printed
, but was obtained in electronic format directly from WIPO
nt/pub/published_pct_sequences

A;

arity   2.8%; Score 8; DB 4;   Length 65;
arity   100.0%; Pred. No. 27;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

GLALA 68
        |||
HLALA 29

dard; protein; 65 AA.

(first entry)

rrow expressed probe encoded protein SEQ ID NO: 38012.

arrow expressed exon; gene expression analysis; probe;
ancer; leukaemia; lymphoma; myeloma.

.2.

2001WO-US000668.

2000US-0180312P.
2000US-0207456P.
2000US-00608408.
2000US-00632366.
2000US-0234687P.
2000US-0236359P.
2000GB-00024263.

ULAR DYNAMICS INC.

zel DK, Chen W, Rank DR;

```
Qy   61 LGLGLALA 68
        ||||||||
Db   22 LGLGLALA 29
```

```
RESULT 41
ABG46737
ID   ABG46737 standard; peptide; 65 AA.
XX
AC   ABG46737;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human peptide encoded by genome-derived single exon probe SEQ ID
XX
KW   Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndr
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease.
XX
OS   Homo sapiens.
XX
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US000665.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used
PT   measure gene expression in human lung samples.
XX
PS   Claim 27; SEQ ID NO 36402; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exc
CC   nucleic acid probes for measuring gene expression in a sample deri
CC   from human lung comprising single exon nucleic acid probes having
CC   12614 nucleic acid sequences mentioned in the specification, or th
CC   complements or the 12387 open reading frames derived from the 1261
CC   probes. Also included are a microarray comprising the novel set of
CC   ; the novel set of probes which hybridise at high stringency to a
CC   acid expressed in the human lung; measuring gene expression in a s
CC   derived from human lung, comprising (a) contacting the array with
CC   collection of detectably labeled nucleic acids derived from human
CC   mRNA, and (b) measuring the label detectably bound to each probe o
CC   array; identifying exons in a eukaryotic genome, comprising (a)
CC   algorithmically predicting at least one exon from genomic sequence
CC   the eukaryote; and (b) detecting specific hybridisation of detecta
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon p
CC   having a fragment identical to the predicted exon, the probe is in
CC   in the above mentioned microarray; assigning exons to a single gen
CC   comprising (a) identifying exons from genomic sequence by the meth
CC   above and (b) measuring the expression of each of the exons in sev
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern o
```

Partial adjacent result (left column):

```
;
   2.8%;  Score 8;  DB 4;  Length 65;
  100.0%;  Pred. No. 27;
nservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

ALA 68
|||
ALA 29
```

```
lard; peptide; 65 AA.

first entry)

ptide, SEQ ID No 38009.

cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
olaemia; coronary heart disease.

LAR DYNAMICS INC.

el DK, Chen W, Rank DR;

98/53.

001WO-US000664.

000US-0180312P.
000US-0207456P.
000US-00608408.
000US-00632366.
000US-0234687P.
000US-0236359P.
000GB-00024263.
```

```
erived single exon nucleic acid probes useful for analyzing
n in human adult liver.

ID NO 38009; 658pp; English.

relates to a single exon nucleic acid probe (SENP) (I) for
n gene expression in a sample derived from human adult
ing one of 13109 defined nucleotide sequences given in the
(or complements/ fragments). The probe hybridises at high
a nucleic acid molecule expressed in the human adult liver.
d for predicting, measuring and displaying gene expression
ived from human adult liver. The genes identified may be
netic liver diseases such as cirrhosis,
inaemia, hyperlipidaemia and hypercholesterolaemia which is
h coronary heart disease. ABG47348-ABG59930 represent human
xon encoded peptides of the invention. Note: The sequence
r this patent does not appear in the printed specification
ed in electronic format directly from WIPO at
ub/published_pct_sequences
;
```

```
   2.8%;  Score 8;  DB 4;  Length 65;
rity  100.0%;  Pred. No. 27;
nservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

CC activity of (II) or to treat disease states involving (II). (II)
CC useful for generating antibodies against it, detecting or quanti
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical
CC of sites expressing (II). (I) and (II) are useful for treating d
CC involving aberrant protein expression or biological activity. Th
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutation
CC responsible for genetic disorders or other traits to assess biod
CC and to produce other types of data and products dependent on DNA
CC amino acid sequences. ABG0010-ABG30377 represent novel human di
CC amino acid sequences of the invention. Note: The sequence data f
CC patent did not appear in the printed specification, but was obta
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 69 AA;

Query Match   2.8%; Score 8; DB 4; Length 69;
Best Local Similarity   100.0%; Pred. No. 29;
Matches   8; Conservative   0; Mismatches   0; Indels   0;

Qy   61 LGLGLALA 68
               ||||||||
Db   16 LGLGLALA 23

RESULT 43
ABG03663
ID   ABG03663 standard; protein; 71 AA.
XX
AC   ABG03663;
XX
DT   13-FEB-2002 (first entry)
XX
DE   Novel human diagnostic protein #3654.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO20017506 7-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT, Liu C, Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS67850.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 20; SEQ ID NO 34022; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypept
CC   sequences. (I) is useful as hybridisation probes, polymerase chai
CC   reaction (PCR) primers, oligomers, and for chromosome and gene ma
CC   and in recombinant production of (II). The polynucleotides are al
CC   in diagnostics as expressed sequence tags for identifying express
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II)
CC   useful for generating antibodies against it, detecting or quantit

---

f the exons in the tissues and/or cell types indicates that
ould be assigned to a single gene; a peptide comprising one
ences, mentioned in the specification, or encoded by the
reading frames (ORF). The probes are used for gene expression
1 for identifying exons in a gene, particularly using human
mRNA and for the study of lung diseases such as asthma, lung
ic obstructive pulmonary disease (COPD), interstitial lung
, familial idiopathic pulmonary fibrosis, neurofibromatosis,
rosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
me, sarcoidosis, pulmonary haemosiderosis, pulmonary
e, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
drome, fibrocystic pulmonary dysplasia, primary ciliary
pulmonary hypertension and hyaline membrane disease. The
ence is a peptide/protein encoded by a single exon probe of
n. Note: The sequence data for this patent did not form part
ed specification, but was obtained in electronic format
n WIPO at ftp.wipo.int/pub/published_pct_sequences

AA;
arity   2.8%; Score 8; DB 5; Length 65;
100.0%; Pred. No. 27;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

LALA 68
||||
LALA 29

dard; protein; 69 AA.

(first entry)

diagnostic protein #3714.

osome mapping; gene mapping; gene therapy; forensic;
nt; medical imaging; diagnostic; genetic disorder.

2.

2001WO-US008631.

2000US-00540217.
2000US-00649167.

INC.

Liu C, Tang YT;

362/73.
910.

polynucleotide and encoded polypeptides, useful in
forensics, gene mapping, identification of mutations
or genetic disorders or other traits and to assess

ID NO 34082; 103pp; English.

relates to isolated polynucleotide (I) and polypeptide (II)
) is useful as hybridisation probes, polymerase chain
) primers, oligomers, and for chromosome and gene mapping,
inant production of (II). The polynucleotides are also used
s as expressed sequence tags for identifying expressed
useful in gene therapy techniques to restore normal

```
n tissue, as molecular weight markers and as a food
II) and its binding partners are useful in medical imaging
essing (II). (I) and (II) are useful for treating disorders
rrant protein expression or biological activity. The
nd polynucleotide sequences have applications in
forensics, gene mapping, identification of mutations
or genetic disorders or other traits to assess biodiversity
e other types of data and products dependent on DNA and
quences. ABG00010-ABG30377 represent novel human diagnostic
quences of the invention. Note: The sequence data for this
t appear in the printed specification, but was obtained in
rmat directly from WIPO at
pub/published_pct_sequences

A;

arity   2.8%;  Score 8;  DB 4;  Length 71;
100.0%;  Pred. No. 29;
onservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

LALA 68
     |||
LALA 23

dard; protein; 84 AA.

(first entry)

ium acnes immunogenic protein #22178.

; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
phthalmitis; bone; joint; central nervous system; ELISA;
esion; acne vulgaris; enzyme linked immunosorbent assay;
; osteopathic; neuroprotectant.

ium acnes.

001WO-US012865.

000US-0199047P.
000US-0208841P.
000US-0216747P.

CORP.

ersing DH, Mitcham JL, Wang SS, Bhatia A;
J, Zhang Y, Jen S, Carter D;

74/71.
15.

ium acnes polypeptides and nucleic acids useful for
ainst and diagnosing infections, especially useful for
vulgaris.

ID NO 22477; 1069pp; English.

9105-AAU68017 represent Propionibacterium acnes immunogenic
The proteins and their associated DNA sequences are used in
prevention and diagnosis of medical conditions caused by
disorders include SAPHO syndrome (synovitis, acne,
pertosis and osteomyelitis), uveitis and endophthalmitis.
so involved in infections of bone, joints and the central
, however it is particularly involved in the inflammatory
```

```
CC  lesions associated with acne vulgaris. A method for detecting the
CC  presence or absence of P. acnes in a patient comprises contacting
CC  sample with a binding agent that binds to the proteins of the inv
CC  and determining the amount of bound protein in the sample. The
CC  polypeptides may be used as antigens in the production of antibod
CC  specific for P. acnes proteins. These antibodies can be used to
CC  downregulate expression and activity of P. acnes polypeptides and
CC  therefore treat P. acnes infections. The antibodies may also be u
CC  diagnostic agents for determining P. acnes presence, for example,
CC  enzyme linked immunosorbent assay (ELISA). Note: The sequence dat
CC  this patent did not form part of the printed specification, but w
CC  obtained in electronic format directly from WIPO at
CC  ftp.wipo.int/pub/published_pct_sequences
XX
SQ  Sequence 84 AA;

Query Match      2.8%;  Score 8;  DB 4;  Length 84;
Best Local Similarity  100.0%;  Pred. No. 34;
Matches   8;  Conservative  0;  Mismatches  0;  Indels  0;  G

Qy  259 LRIRTLPW 266
        ||||||||
Db   34 LRIRTLPW 41

RESULT 45
ABM57801
ID  ABM57801 standard; protein; 84 AA.
XX
AC  ABM57801;
XX
DT  20-OCT-2003  (first entry)
XX
DE  Propionibacterium acnes predicted ORF-encoded polypeptide #22477.
XX
KW  Acne vulgaris; antiseborrhoeic; dermatological; antibacterial;
KW  immunostimulant; immune response; vaccine.
XX
OS  Propionibacterium acnes.
XX
PN  WO200303515-A1.
XX
PD  24-APR-2003.
XX
PF  11-OCT-2002; 2002WO-US032727.
XX
PR  15-OCT-2001; 2001US-00978825.
XX
PA  (CORI-) CORIXA CORP.
XX
PI  Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL,
PI  Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter
PI  Barth B, Vallieve-Douglass J;
XX
DR  WPI; 2003-381789/36.
DR  N-PSDB; ACF64544.
XX
PT  New Propionibacterium acnes polypeptides and polynucleotides encod
PT  polypeptide, useful for diagnosing, preventing or treating acne vu
PT  or for stimulating an immune response specific for a P. acnes prot
XX
PS  Example 1; SEQ ID NO 22477; 1481pp; English.
XX
CC  The invention relates to an isolated polynucleotide (ACF64435-ACF6
CC  encoding a Propionibacterium acnes protein. The invention also rel
CC  polypeptides encoded by the polynucleotides (ABM35624-ABM64536) an
CC  immunogenic fragments of P. acnes polypeptides. The invention
CC  additionally encompasses expression vectors and host cells compris
CC  polynucleotide of the invention; antibodies against polypeptides c
CC  invention; fusion proteins comprising a polypeptide of the inventi
CC  method for stimulating an immune response specific for a P. acnes
CC  polypeptide and an isolated T cell population comprising T cells p
CC  via this method, a vaccine composition (comprising P. acnes polype
```

```
CC   activity of (II) or to treat disease states involving (II). (II)
CC   useful for generating antibodies against it, detecting or quanti
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical
CC   of sites expressing (III). (I) and (II) are useful for treating d
CC   involving aberrant protein expression or biological activity. Th
CC   polypeptide and polynucleotide sequences have application in
CC   diagnostics, forensics, gene mapping, identification of mutation
CC   responsible for genetic disorders or other traits to assess biod
CC   and to produce other types of data and products dependent on DNA
CC   amino acid sequences. ABG00010-ABG30377 represent novel human di
CC   amino acid sequences of the invention. Note: The sequence data f
CC   patent did not appear in the printed specification, but was obta
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 110 AA;

Query Match            2.8%;  Score 8;  DB 4;  Length 110;
Best Local Similarity 100.0%;  Pred. No. 45;
Matches    8;  Conservative   0;  Mismatches   0;  Indels   0;

Qy    61 LGLGLALA 68
         ||||||||
Db    44 LGLGLALA 51


RESULT 47
AAO08094
ID   AAO08094 standard; protein; 117 AA.
XX
AC   AAO08094;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human polypeptide SEQ ID NO 21986.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; gene
KW   vaccine; peptide therapy; stem cell growth factor; haematopoiesi
KW   tissue growth factor; immunomodulatory; cancer; leukaemia;
KW   nervous system disorders; arthritis; inflammation.
XX
OS   Homo sapiens.
XX
PN   WO200164835-A2.
XX
PD   07-SEP-2001.
XX
PF   26-FEB-2001; 2001WO-US004927.
XX
PR   28-FEB-2000; 2000US-00515126.
PR   18-MAY-2000; 2000US-00577409.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT, Liu C, Drmanac RT;
XX
DR   WPI; 2001-514838/56.
DR   N-PSDB; AAI88025.
XX
PT   Isolated nucleic acids and polypeptides, useful for preventing d:
PT   and treating e.g. leukemia, inflammation and immune disorders.
XX
PS   Claim 20; SEQ ID NO 21986; 1399pp + Sequence Listing; English.
XX
CC   The invention relates to human polynucleotides (AAI79941-AAI93841
CC   the encoded proteins (AAO00010-AAO13910) that exhibit activity el
CC   cytokine, cell proliferation or cell differentiation or which ma}
CC   production of other cytokines in other cell population. The
CC   polynucleotides and polypeptides are useful in gene therapy, vacc
CC   peptide therapy. The polypeptides have various cytokine-like acti
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activit
```

```
des, antibodies, fusion proteins, T cell populations, or
enting cells that express the polypeptide]; a method and kit
g or determining the presence or absence of P. acnes in a
a method for inhibiting the development of P. acnes in a
P. acnes polypeptides, polynucleotides, antibodies, fusion
cell populations or antigen-presenting cells that express the
are useful for diagnosing, preventing or treating acne
for stimulating an immune response specific for a P. acnes
polynucleotides can also be used as probes or primers for
hybridisation. The vaccine composition is useful for the
of an immune response against P. acnes, or for treating acne,
is useful for performing a diagnostic assay. The present
resents a polypeptide predicted to be encoded by an ORF (open
e) contained within the P. acnes polynucleotides of the
ote: The sequence data for this patent did not form part of
specification, but was obtained in electronic format directly
ftp.wipo.int/pub/published_pct_sequences

AA;

larity       2.8%;  Score 8;  DB 6;  Length 84;
          100.0%;  Pred. No. 34;
Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

RTLPW 266
||||||
RTLPW 41


ddard; protein; 110 AA.

(first entry)

diagnostic protein #20250.

osome mapping; gene mapping; gene therapy; forensic;
ent; medical imaging; diagnostic; genetic disorder.

A2.

2001WO-US008631.

2000US-00540217.
2000US-00649167.

] INC.

Liu C, Tang YT;

362/73.
446.

polynucleotide and encoded polypeptides, useful in
forensics, gene mapping, identification of mutations
for genetic disorders or other traits and to assess

) ID NO 50618; 103pp; English.

n relates to isolated polynucleotide (I) and polypeptide (II)
) is useful as hybridisation probes, polymerase chain
) primers, oligomers, and for chromosome and gene mapping,
inant production of (II). The polynucleotides are also used
s as expressed sequence tags for identifying expressed
useful in gene therapy techniques to restore normal
```

Db   16 LLLAVVSL 23

RESULT 49
ABP28041
ID   ABP28041 standard; protein; 190 AA.
XX
AC   ABP28041;
XX
DT   02-JUL-2002 (first entry)
XX
DE   Streptococcus polypeptide SEQ ID NO 5258.
XX
KW   Streptococcus; GAS; GBS; group B streptococcus; Streptococcus aga
KW   group A streptococcus; Streptococcus pyogenes; antibacterial;
KW   antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS   Streptococcus agalactiae.
XX
PN   WO200234771-A2.
XX
PD   02-MAY-2002.
XX
PF   29-OCT-2001; 2001WO-GB004789.
XX
PR   27-OCT-2000; 2000GB-00026333.
PR   24-NOV-2000; 2000GB-00028727.
PR   07-MAR-2001; 2001GB-00005640.
XX
PA   (CHIR-) CHIRON SPA.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Telford J, Masignani V, Margarit Y RosI, Grandi G, Fraser C;
PI   Tettelin H;
XX
DR   WPI; 2002-352536/38.
DR   N-PSDB; ABN68672.
XX
PT   New Streptococcus protein for the treatment or prevention of infe
PT   disease caused by Streptococcus bacteria, such as meningitis, and
PT   detecting a compound that binds to the protein.
XX
PS   Claim 1; Page 3689; 4525pp; English.
XX
CC   The invention relates to a protein (ABP25413-ABP30895) from group
CC   streptococcus/GBS (Streptococcus agalactiae) or group A streptoco
CC   (Streptococcus pyogenes), comprising one of 5483 sequences (S1),
CC   the specification. The proteins have antibacterial and antiinflam
CC   activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC   antibodies that bind (I) are used in the manufacture of medicamen
CC   the treatment or prevention of infection or disease caused by
CC   Streptococcus bacteria, particularly S. agalactiae and S. pyrogen
CC   Nucleic acids encoding (I) are used to detect Streptococcus in a
CC   biological sample. (I) is used to determine whether a compound bi
CC   (I). A composition comprising (I) or a nucleic acid encoding (I),
CC   used as a vaccine or diagnostic composition. The disease caused b
CC   Streptococcus that is prevented or treated may be meningitis. Nuc
CC   acid encoding (I) may be used to recombinantly produce (I) and ma
CC   used in gene therapy. Antibodies to (I) are used for affinity
CC   chromatography, immunoassays, and distinguishing/identifying
CC   Streptococcus proteins
XX
SQ   Sequence 190 AA;

Query Match   2.8%;   Score 8; DB 5; Length 190;
Best Local Similarity   100.0%;   Pred. No. 75;
Matches   8;   Conservative   0;   Mismatches   0;   Indels   0;   Ga

Qy   258 SLRIRTLP 265
         ||||||||
Db   50 SLRIRTLP 57

---

in activity and may be useful in the diagnosis and/or
cancer, leukaemia, nervous system disorders, arthritis and
Note: The sequence data for this patent did not form part
i specification, but was obtained in electronic format
WIPO at ftp.wipo.int/pub/published_pct_sequences

AA;

arity   2.8%;   Score 8; DB 4;   Length 117;
      100.0%;   Pred. No. 47;
onservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

ALA 68
|||
ALA 61

dard; protein; 184 AA.

(first entry)

lanogaster polypeptide SEQ ID NO 29928.

evelopmental biology; cell signalling; insecticide;

langoaster.

2.

2001WO-US009231.

2000US-0191637P.
2000US-00614150.

ip NY.

iams M, Li PWD, Myers EW;

860/75.
315.

nucleic acid detection reagent for detecting 1000 or more
ecting 1000 or more genes from Drosophila. The invention is
lopmental biology in elucidating cell signalling and
osophila and for elucidating cell signalling and cell-cell

SQ ID NO 29928; 21pp + Sequence Listing; English.

relates to an isolated nucleic acid detection reagent
ecting 1000 or more genes from Drosophila. The invention is
lopmental biology in elucidating cell signalling and
ractions in higher eukaryotes for the development of
therapeutics and pharmaceutical drugs. The invention
mic DNA sequences (ABL16176-ABL30511), expressed DNA
01840-ABL16175) and the encoded proteins (ABB5737-
sequence data for this patent did not form part of the
ication, but was obtained in electronic format directly
ftp.wipo.int/pub/published_pct_sequences

AA;

arity   2.8%;   Score 8; DB 4;   Length 184;
      100.0%;   Pred.No. 73;
onservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

VSL 79
|||

KW   infertility; pregnancy disorder; anovulation; polycystic ovary s
KW   PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection
KW   inflammatory condition; immune disorder; blood disorder;
KW   cardiovascular condition; respiratory disorder; neurological diso
KW   gastrointestinal disorder; urinary system disorder; drug screeni
KW   gene therapy; chromosome mapping; forensic analysis;
KW   antibody preparation; cytostatic; immunomodulatory; neuroprotect
KW   antiinflammatory; gynaecological; reproductive.
OS   Homo sapiens.
XX
PN   WO20020677-A1.
XX
PD   03-JAN-2002.
XX
PF   07-JUN-2001; 2001WO-US018569.
XX
PR   07-JUN-2000; 2000US-0209467P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Birse CE, Rosen CA;
XX
DR   WPI; 2002-147878/19.
DR   N-PSDB; ABQ54751.
XX
PT   Isolated nucleic acid molecules encoding novel ovarian polypepti
PT   useful in the prevention, treatment and diagnosis of cancer (e.g
PT   cancer), immune disorders, cardiovascular disorders and neurolog
PT   diseases.
XX
PS   Claim 11; SEQ ID NO 2806; 2922pp; English.
XX
CC   The invention relates to 2175 novel human ovarian antigens (ABP4
CC   ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and al
CC   encompasses polypeptides 90% identical and polynucleotides 95% i
CC   to the sequences of the invention. The invention additionally re
CC   recombinant vectors and host cells comprising human ovarian anti
CC   polynucleotides, antibodies against human ovarian antigens, and
CC   of ovarian antigen polynucleotides and polypeptides in diagnosin
CC   treating, prognosing or preventing various ovary and/or breast-r
CC   disorders. Such conditions include ovarian cancer and breast can
CC   metastatic tumours of ovarian or breast origin, reproductive sys
CC   disorders (e.g., infertility, disorders of pregnancy, anovulatio
CC   polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), en
CC   disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and
CC   shock syndrome), inflammatory conditions (e.g., mastitis, oophor
CC   vaginitis), immune disorders (e.g., congenital and acquired
CC   immunodeficiencies, autoimmune oophoritis, systemic lupus erythe
CC   blood-related disorders (e.g., anaemia), cardiovascular disorder
CC   respiratory disorders, neurological disorders, gastrointestinal
CC   and urinary system disorders. Ovarian antigen polypeptides and
CC   polynucleotides may also be used in screening for compounds whic
CC   modulate ovarian antigen expression or activity. The polynucleoti
CC   further be used for gene therapy, chromosome mapping, in the
CC   identification of individuals and in forensic analysis, and the
CC   polypeptides may be used as food additives or to prepare antibodi
CC   useful in disease diagnosis, drug targeting and phenotyping. The
CC   sequence represents a human ovarian antigen of the invention. Not
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly fr
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 222 AA;

Query Match   2.8%;   Score 8;   DB 5;   Length 222;
Best Local Similarity   100.0%;   Pred. No. 87;
Matches   8;   Conservative   0;   Mismatches   0;   Indels   0;   C

Qy   60 ALGLGLAL 67
         ||||||||
Db   190 ALGLGLAL 197

---

dard; protein; 198 AA.

(first entry)

coded protein SEQ ID NO: 1209.

pig; cow; fruit fly; yeast; hamster; macaque; horse;
ey; dog; sea urchin; expressed sequence tag; EST;
forensic test; gene mapping; genetic disorder; biodiversity;
nutrition.

2001WO-US002687.

2000US-00491404.
2000US-00617746.
2000US-00631451.
2000US-00663870.

INC.

C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
ac RA, Zhang J, Werhman T;

164/51.
343.

peptide for treatment of diseases, diagnostics, raising
d research use.

e 873; 1275pp; English.

nvention provides the protein and coding sequences of novel
a variety of organisms, including human, dog, cat, horse,
ster, monkey, macaque, yeast, bacteria, fruit fly, sea
mato. These were derived from expressed sequence tags (ESTs)
nism of interest. They can be used in diagnostics,
ne mapping, identification of mutations, to assess
and for nutritional purposes. The present sequence is a
e invention

AA;

arity   2.8%;   Score 8;   DB 4;   Length 198;
onservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   PLAL 61
        |||
Db   PLAL 9

dard; protein; 222 AA.

(first entry)

antigen HPAMG11, SEQ ID NO:2806.

n antigen; ovary; ovarian; breast; cancer; tumour;
er; breast cancer; tumour; reproductive system disorder;

ABG01186
ID   ABG01186 standard; protein; 307 AA.
XX
AC   ABG01186;
XX
DT   13-FEB-2002 (first entry)
XX
DE   Novel human diagnostic protein #1177.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT, Liu C, Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS65373.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 20; SEQ ID NO 31545; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypept
CC   sequences. (I) is useful as hybridisation probes, polymerase chai
CC   reaction (PCR) primers, oligomers, and for chromosome and gene ma
CC   and in recombinant production of (II). The polynucleotides are al
CC   in diagnostics as expressed sequence tags for identifying express
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II)
CC   useful for generating antibodies against it, detecting or quantit
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical i
CC   of sites expressing (II). (I) and (II) are useful for treating di
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodi
CC   and to produce other types of data and products dependent on DNA
CC   amino acid sequences. ABG00010-ABG30377 represent novel human dia
CC   amino acid sequences of the invention. Note: The sequence data fo
CC   patent did not appear in the printed specification, but was obtai
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 307 AA;

Query Match   2.8%; Score 8; DB 4; Length 307;
Best Local Similarity   100.0%;
Matches   8; Conservative   0; Mismatches   0; Indels   0; G

Qy   77 VSLGSRAS 84
              ||||||||
Db   3 VSLGSRAS 10

RESULT 54
ADB79952
ID   ADB79952 standard; protein; 342 AA.

PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT, Liu C, Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS77578.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutation
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 20; SEQ ID NO 43750; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypep
CC   sequences. (I) is useful as hybridisation probes, polymerase cha
CC   reaction (PCR) primers, oligomers, and for chromosome and gene m
CC   and in recombinant production of (II). The polynucleotides are a
CC   in diagnostics as expressed sequence tags for identifying expres
CC   genes. (I) is useful in gene therapy techniques to restore norma
CC   activity of (II) or to treat disease states involving (II) . (II)
CC   useful for generating antibodies against it, detecting or quanti
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical
CC   of sites expressing (II). (I) and (II) are useful for treating d
CC   involving aberrant protein expression or biological activity. Th
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutation
CC   responsible for genetic disorders or other traits to assess biod
CC   and to produce other types of data and products dependent on DNA
CC   amino acid sequences. ABG00010-ABG30377 represent novel human di
CC   amino acid sequences of the invention. Note: The sequence data f
CC   patent did not appear in the printed specification, but was obta
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 370 AA;

Query Match   2.8%; Score 8; DB 4; Length 370;
Best Local Similarity   100.0%;   Pred. No. 1.4e+02;
Matches   8; Conservative   0; Mismatches   0; Indels   0;

Qy   63 LGLALACL 70
        ||||||||
Db   28 LGLALACL 35

RESULT 56
ABG05012
ID   ABG05012 standard; protein; 370 AA.
XX
AC   ABG05012;
XX
DT   13-FEB-2002 (first entry)
XX
DE   Novel human diagnostic protein #5003.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX

(first entry)

BB 1 progression enhanced protein, SEQ ID 192.

ain; streptozocin-induced diabetes; rat.

gicus.

2002EP-00255249.

2001GB-00018354.
2002GB-00002910.

ER LAMBERT CO.

, Dixon AK, Lee K, Pinnock RD;

5407/38.
9953.

ed gene sequences and encoded polypeptides that are
n the spinal cord in response to streptozocin-induced
screening compounds for the treatment of pain, or for
ain.

a 326-327; 334pp; English.

invention relates to nucleotide sequences which are useful in
of compounds for the treatment of pain, or for the
g of compounds for the treatment of pain. The nucleotide sequences are up-regulated in the
in response to streptozocin-induced diabetes. The present
used to illustrate the invention.

AA;

arity   2.8%; Score 8; DB 7; Length 342;
100.0%; Pred. No. 1.3e+02;
onservative   0; Mismatches   0; Indels   0; Gaps   0;

GLAL 67
|||
GLAL 317

dard; protein; 370 AA.

(first entry)

diagnostic protein #13382.

some mapping; gene mapping; gene therapy; forensic;
nt; medical imaging; diagnostic; genetic disorder.

2.

2001WO-US008631.
2000US-00540217.

PI   Drmanac RT, Liu C, Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS82302.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 20; SEQ ID NO 48474; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypept
CC   sequences. (I) is useful as hybridisation probes, polymerase chai
CC   reaction (PCR) primers, oligomers, and for chromosome and gene ma
CC   and in recombinant production of (II). The polynucleotides are al
CC   in diagnostics as expressed sequence tags for identifying express
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II)
CC   useful for generating antibodies against it, detecting or quantit
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical i
CC   of sites expressing (II). (I) and (II) are useful for treating di
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodi
CC   and to produce other types of data and products dependent on DNA.
CC   amino acid sequences. ABG00010-ABG30377 represent novel human dia
CC   amino acid sequences of the invention. Note: The sequence data fo
CC   patent did not appear in the printed specification, but was obtai
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 370 AA;

Query Match   2.8%; Score 8; DB 4; Length 370;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches   8; Conservative   0; Mismatches   0; Indels   0; Gi

Qy   63 LGLALACL 70
              ||||||||
Db   28 LGLALACL 35

RESULT 58
ABU36537
ID   ABU36537 standard; protein; 372 AA.
XX
AC   ABU36537;
XX
DT   19-JUN-2003 (first entry)
XX
DE   Protein encoded by Prokaryotic essential gene #22064.
XX
KW   Antisense; prokaryotic essential gene; cell proliferation; drug de
XX
OS   Mycobacterium tuberculosis.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-0094899.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-0007285.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX

INC.

iu C, Tang YT;

362/73.
199.

polynucleotide and encoded polypeptides, useful in
forensics, gene mapping, identification of mutations
or genetic disorders or other traits and to assess

ID NO 35371; 103pp; English.

relates to isolated polynucleotide (I) and polypeptide (II)
is useful as hybridisation probes, polymerase chain
primers, oligomers, and for chromosome and gene mapping,
nant production of (II). The polynucleotides are also used
s as expressed sequence tags for identifying expressed
useful in gene therapy techniques to restore normal
II) or to treat disease states involving (II). (II) is
erating antibodies against it, detecting or quantitating a
n tissue, as molecular weight markers and as a food
II) and its binding partners are useful in medical imaging
ssing (II). (I) and (II) are useful for treating disorders
rant protein expression or biological activity. The
nd polynucleotide sequences have applications in
orensics, gene mapping, identification of mutations
r genetic disorders or other traits to assess biodiversity
other types of data and products dependent on DNA and
quences. ABG00010-ABG30377 represent novel human diagnostic
quences of the invention. Note: The sequence data for this
appear in the printed specification, but was obtained in
mat directly from WIPO at
pub/published_pct_sequences

A;

rity   2.8%; Score 8; DB 4; Length 370;
       100.0%; Pred. No. 1.4e+02;
nservative   0; Mismatches   0; Indels   0; Gaps   0;

ACL 70
|||
ACL 35

ard; protein; 370 AA.

first entry)

agnostic protein #18106.

ome mapping; gene mapping; gene therapy; forensic;
c; medical imaging; diagnostic; genetic disorder.

001WO-US008631.

000US-0540217.
000US-0649167.

INC.

dio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
vick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
9926/02.
0407.
XX
a nucleic acids, useful for identifying proteins or screening
as nucleic acids required for cellular proliferation to
idate molecules for rational drug discovery programs.
XX
O ID NO 64461; 1766pp; English.
XX
l relates to an isolated nucleic acid comprising any one of
sense sequences given in the specification where expression
c acid inhibits proliferation of a cell. Also included are:
comprising a promoter operably linked to the nucleic acid
olypeptide whose expression is inhibited by the antisense
(2) a host cell containing the vector; (3) an isolated
or its fragment whose expression is inhibited by the
cleic acid; (4) an antibody capable of specifically binding
de; (5) producing the polypeptide; (6) inhibiting cellular
or the activity of a gene in an operon required for
; (7) identifying a compound that influences the activity of
uct or that has an activity against a biological pathway
proliferation, or that inhibits cellular proliferation; (8)
gene required for cellular proliferation or the biological
nich a proliferation-required gene or its gene product lies
which the test compound that inhibits proliferation of an
; (9) manufacturing an antibiotic; (10) profiling a
tivity; (11) a culture comprising strains in which the gene
rexpressed or underexpressed; (12) determining the extent
of the strains is present in a culture or collection of
13) identifying the target of a compound that inhibits the
l of an organism. The antisense nucleic acids are useful for
proteins or screening for homologous nucleic acids required
proliferation to isolate candidate molecules for rational
y programs, or for screening homologous nucleic acids
proliferation in cells other than S. aureus, S. typhimurium,
or P. aeruginosa. The present sequence is encoded by one of
okaryotic essential genes. Note: The sequence data for this
t form part of the printed specification, but was obtained
format directly from WIPO at
pub/published_pct_sequences
XX
AA;

2.8%; Score 8; DB 6; Length 372;
arity 100.0%; Pred. No. 1.4e+02;
onservative 0; Mismatches 0; Indels 0; Gaps 0;

LALA 68
|||
LALA 103

dard; protein; 424 AA.

(first entry)

diagnostic protein #15604.

some mapping; gene mapping; gene therapy; forensic;
nt; medical imaging; diagnostic; genetic disorder.

2.

XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS79800.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutation
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 45972; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypept
CC sequences. (I) is useful as hybridisation probes, polymerase cha
CC reaction (PCR) primers, oligomers, and for chromosome and gene ma
CC and in recombinant production of (II). The polynucleotides are a
CC in diagnostics as expressed sequence tags for identifying express
CC genes. (I) is useful in gene therapy techniques to restore norma
CC activity of (II) or to treat disease states involving (II). (II)
CC useful for generating antibodies against it, detecting or quanti
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical
CC of sites expressing (II). (I) and (II) are useful for treating d
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biod
CC and to produce other types of data and products dependent on DNA
CC amino acid sequences. ABG00010-ABG30377 represent novel human di
CC amino acid sequences of the invention. Note: The sequence data fo
CC patent did not appear in the printed specification, but was obtai
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 424 AA;

Query Match   2.8%; Score 8; DB 4; Length 424;
Best Local Similarity   100.0%; Pred. No. 1.6e+02;
Matches   8; Conservative   0; Mismatches   0; Indels   0; G

Qy   60 ALGLGLAL 67
         ||||||||
Db   392 ALGLGLAL 399

RESULT 60
AAM23752
ID AAM23752 standard; protein; 430 AA.
XX
AC AAM23752;
XX
DT 12-OCT-2001 (first entry)
XX
DE Human EST encoded protein SEQ ID NO: 1277.
XX
KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW diagnostics; forensic test; gene mapping; genetic disorder; biodi
KW gene therapy; nutrition.
XX
OS Homo sapiens.
XX
PN WO20015477-A2.
XX
PD 02-AUG-2001.

XX
DR   2001WO-US002687.
DR
2000US-00491404.
2000US-00617746.
2000US-00631451.
2000US-00663870.
XX
INC.
XX
C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
ac RA, Zhang J, Werhman T;
XX
164/51.
411.
XX
peptide for treatment of diseases, diagnostics, raising
d research use.
XX
a 920; 1275pp; English.
XX
invention provides the protein and coding sequences of novel
a variety of organisms, including human, dog, cat, horse,
ster, monkey, macaque, yeast, bacteria, fruit fly, sea
nato. These were derived from expressed sequence tags (ESTs)
nism of interest. They can be used in diagnostics,
ne mapping, identification of mutations, to assess
and for nutritional purposes. The present sequence is a
e invention
XX
AA;
rity   2.8%;   Score 8;   DB 4;   Length 430;
100.0%;   Pred. No. 1.6e+02;
onservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

LAL 61
|||
LAL 9

lard; protein; 431 AA.
first entry)
equence Seq ID508 related to grain filling.
ology; carbohydrate synthesis; carbohydrate metabolism;
egradation; carbohydrate; plant grain; grain filling; corn;
; canola; cotton; peanut; sorghum; tobacco; sugarbeet;
rotein; oil; starch; fibre; moisture content; cereal grain;
t.
TA PARTICIPATIONS AG.
XX
W, Briggs S, Cooper B, Goff SA, Moughamer T;
Katagiri F, Kreps J, Provart N, Ricke D;
XX
002WO-IB002450.
XX
001US-0300112P.
001US-0325277P.
001US-0342327P.
XX
2.

XX
DR   WPI; 2003-229341/22.
DR   N-PSDB; ADC08202.
XX
PT   New plant genes encoding polypeptides having an activity involved
associated with the synthesis, metabolism or degradation of carb
in the plant grain useful in generating plants having improved
nutritional properties.
XX
PS   Claim 34; SEQ ID NO 508; 130pp; English.
XX
CC   This invention, in the area of plant biotechnology, relates to n
CC   polynucleotides comprising a nucleotide sequence encoding a prote
CC   is involved in or associated with the synthesis, metabolism or
CC   degradation of carbohydrates in the plant grain and the expressio
CC   which is up-regulated during grain filling. The plant is selected
CC   corn, tomato, banana, canola, cotton, peanut, sorghum, tobacco,
CC   sugarbeet, wheat, and rice. The invention may be useful for the
CC   improvement of protein, oil, starch, fibre and moisture content c
CC   cereal grains. In addition, carbohydrate levels may be modified t
CC   desirable level using the present invention. The present sequence
CC   amino acid sequence of a rice protein of the invention. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly fro
CC   at ftp.wipo.int/pub/publishedpct_sequences.
XX
SQ   Sequence 431 AA;

Query Match   2.8%;   Score 8;   DB 7;   Length 431;
Best Local Similarity   100.0%;   Pred. No. 1.6e+02;
Matches   8;   Conservative   0;   Mismatches   0;   Indels   0;   G

Qy   42 QRRRGRRG 49
        ||||||||
Db 372 QRRRGRRG 379

RESULT 62
ADC64562
ID   ADC64562 standard; protein; 431 AA.
XX
AC   ADC64562;
XX
DT   01-JAN-2004 (first entry)
XX
DE   Synechococcus sp. Synwh0268 protein.
XX
KW   Plant growth; commercial yield; plant breeding; fruit yield;
KW   flowering rate; Synwh0268.
XX
OS   Synechococcus sp.; WH 8102.
XX
PN   US2003192076-A1.
XX
PD   09-OCT-2003.
XX
PF   10-APR-2003; 2003US-00410432.
XX
PR   26-MAR-2002; 2002WO-IL000250.
XX
PA   (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
PI   Kaplan A, Lieman-Hurwitz J, Schatz D, Mittler R, Rachmilevitch
XX
DR   WPI; 2003-831832/77.
XX
PT   Obtaining plants having enhanced growth and/or fruit yield and/or
PT   flowering rate, specifically C3 plants grown under limiting condit
PT   useful in plant molecular biology and commercial plant breeding.
XX
PS   Claim 2; Fig 11; 47pp; English.
XX
CC   The present invention relates to a method of obtaining plants with

Matches   8; Conservative   0; Mismatches   0; Indels   0;

Qy   56 LVPLALGL 63
         ||||||||
Db   170 LVPLALGL 177

RESULT 64
ADA54710
ID   ADA54710 standard; protein; 472 AA.
XX
AC   ADA54710;
XX
DT   20-NOV-2003 (first entry)
XX
DE   Human protein, SEQ ID 2278.
XX
KW   Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Noc
KW   Gene Therapy; human; secretory protein; membrane proteins; cance
KW   inflammatory disease; osteoporosis; neurological disease.
XX
OS   Homo sapiens.
XX
PN   EP1293569-A2.
XX
PD   19-MAR-2003.
XX
PF   21-MAR-2002; 2002EP-00006586.
XX
PR   14-SEP-2001; 2001JP-00328381.
PR   24-JAN-2002; 2002US-0350435P.
XX
PA   (HELI-) HELIX RES INST.
PA   (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI   Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii
PI   Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tam
PI   Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX
DR   WPI; 2003-395539/38.
DR   N-PSDB; ADA53071.
XX
PT   New polynucleotides encoding full-length polypeptides, e.g. secr
PT   and/or membrane proteins, useful for developing medicines for di
PT   which the gene is involved, or as target molecules for gene ther
XX
PS   Claim 14; SEQ ID NO 2278; 205pp; English.
XX
CC   The present invention relates to novel human secretory or membra
CC   proteins (ADA54072-ADA55710) and their coding sequences (ADA5243.
CC   ADA54071). The coding sequences are useful in the gene therapy o
CC   diseases caused by abnormalities of the proteins, e.g. cancer,
CC   inflammatory diseases, osteoporosis or neurological disease.
XX
SQ   Sequence 472 AA;

Query Match   2.8%; Score 8; DB 6; Length 472;
Best Local Similarity 100.0%; Pred. No.1.8e+02;
Matches   8; Conservative   0; Mismatches   0; Indels   0; (

Qy   60 ALGLGLAL 67
         ||||||||
Db   440 ALGLGLAL 447

RESULT 65
ABG20260
ID   ABG20260 standard; protein; 586 AA.
XX
AC   ABG20260;
XX
DT   18-FEB-2002 (first entry)
XX

---

wth and/or commercial yield under growth limiting conditions.
omprises obtaining a population of plants transformed to
lypeptide having at least 60% sequence identity to any of 8
d sequences, growing the plants and selecting plants
he polypeptide. The methods and compositions of the present
e useful in commercial plant breeding, particularly for
lants having enhanced growth and/or fruit yield and/or
te. The present sequence represents Synechococcus sp.
otein.

AA;

larity   2.8%; Score 8; DB 7; Length 431;
Conservative   0; Mismatches   0; Indels   0; Gaps   0;

LGLLL 74
|||||
LGLLL 371

ndard; protein; 454 AA.

(first entry)

elanogaster polypeptide SEQ ID NO 516.

developmental biology; cell signalling; insecticide;
al.

elanogaster.

A2.

2001WO-US009231.

200OUS-0191637P.
200OUS-0061415O.

DRP NY.

Adams M, Li PWD, Myers EW;

5860/75.
2011.

nucleic acid detection reagent for detecting 1000 or more
rosophila and for elucidating cell signaling and cell-cell

SEQ ID NO 516; 21pp + Sequence Listing; English.

l relates to an isolated nucleic acid detection reagent
etecting 1000 or more genes from Drosophila. The invention is
velopmental biology and in elucidating cell signalling and
eractions in higher eukaryotes for the development of
therapeutics and pharmaceutical drugs. The invention
omic DNA sequences (ABL6176-ABL30511), expressed DNA
3L01840-ABL16175) and the encoded proteins (ABB57737-
he sequence data for this patent did not form part of the
fication, but was obtained in electronic format directly
ftp.wipo.int/pub/published_pct_sequences

AA;

arity   2.8%; Score 8; DB 4; Length 454;
arity 100.0%; Pred. No.1.7e+02;

KW   Human; pain; neuronal tissue; gene therapy;
KW   spinal segmental nerve injury; chronic constriction injury; CCI;
KW   spared nerve injury; SNI; Chung.
XX
OS   Homo sapiens.
XX
PN   WO2003016475-A2.
XX
PD   27-FEB-2003.
XX
PF   14-AUG-2002; 2002WO-US025765.
XX
PR   14-AUG-2001; 2001US-0312147P.
PR   01-NOV-2001; 2001US-0346382P.
PR   26-NOV-2001; 2001US-0333347P.
XX
PA   (GEHO) GEN HOSPITAL CORP.
PA   (FARB) BAYER AG.
XX
PI   Woolf C, D'urso D, Befort K, Costigan M;
XX
DR   WPI; 2003-268312/26.
DR   GENBANK; P36269.
XX
PT   New composition comprising two or more isolated polypeptides, use
PT   preparing a medicament for treating pain in an animal.
XX
PS   Claim 1; Page; 1017pp; English.
XX
CC   The invention discloses a composition comprising two or more isol
CC   or human polynucleotides or a polynucleotide which represents a f
CC   derivative or allelic variation of the nucleic acid sequence. Als
CC   claimed are a vector comprising the novel polynucleotide, a host
CC   comprising the vector, a method for identifying a nucleotide sequ
CC   which is differentially regulated in an animal subjected to pain
CC   kit to perform the method, an array, a method for identifying an
CC   that increases or decreases the expression of the polynucleotide
CC   that is differentially expressed in neuronal tissue of a first ar
CC   subjected to pain, a method for identifying a compound which regu
CC   the expression of a polynucleotide sequence which is differential
CC   expressed in an animal subjected to pain, a method for identifyin
CC   compound that regulates the activity of one or more of the
CC   polynucleotides, a method for producing a pharmaceutical composit
CC   method for identifying a compound or small molecule that regulate
CC   activity in an animal of one or more of the polypeptides given in
CC   specification, a method for identifying a compound useful in trea
CC   pain and a pharmaceutical composition comprising the one or more
CC   polypeptides or their antibodies. The polynucleotide or the compo
CC   modulates its activity is useful for preparing a medicament for t
CC   pain (e.g. spinal segmental nerve injury (Chung), chronic constri
CC   injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. ge
CC   therapy). The sequence presented is a human protein (shown in Tab
CC   the specification) which is differentially expressed during pain.
CC   The sequence data for this patent did not form part of the printe
CC   specification, but was obtained in electronic form directly from
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 586 AA;

Query Match   2.8%; Score 8; DB 7; Length 586;
Best Local Similarity   100.0%; Pred. No. 2.2e-02;
Matches   8; Conservative   0; Mismatches   0; Indels   0; G

Qy   61 LGLGLALA 68
           ||||||||
Db   14 LGLGLALA 21

RESULT 67
ADE62980
ID   ADE62980 standard; protein; 586 AA.
XX
AC   ADE62980;

---

iagnostic protein #2025I.

...some mapping; gene mapping; gene therapy; forensic;
...nt; medical imaging; diagnostic; genetic disorder.

...2.

...2001WO-US008631.

...2000US-00540217.
...2000US-00649167.

... INC.

Liu C, Tang YT;

...362/73.
...447.

...polynucleotide and encoded polypeptides, useful in
...forensics, gene mapping, identification of mutations
...or genetic disorders or other traits and to assess

ID NO 50619; 103pp; English.

...relates to isolated polynucleotide (I) and polypeptide (II)
...) is useful as hybridisation probes, polymerase chain
...) primers, oligomers, and for chromosome and gene mapping,
...inant production of (II). The polynucleotides are also used
...s as expressed sequence tags for identifying expressed
...useful in gene therapy techniques to restore normal
...II) or to treat disease states involving (II). (II) is
...nerating antibodies against it, detecting or quantitating a
...n tissue, as molecular weight markers and as a food
...II) and its binding partners are useful in medical imaging
...essing (II). (I) and (II) are useful for treating disorders
...rrant protein expression or biological activity. The
...nd polynucleotide sequences have application in
...forensics, gene mapping, identification of mutations
...or genetic disorders or other traits to assess biodiversity
...s other types of data and products dependent on DNA and
...quences. ABG00010-ABG30377 represent novel human diagnostic
...quences of the invention. Note: The sequence data for this
...appear in the printed specification, but was obtained in
...rmat directly from WIPO at
...pub/published_pct_sequences

...AA;

...arity   2.8%; Score 8; DB 4; Length 586;
...arity 100.0%; Pred. No. 2.2e+02;
...onservative   0; Mismatches   0; Indels   0; Gaps   0;

...ALA 68
...||||
...ALA 21

...dard; protein; 586 AA.

...(first entry)

...P36269, SEQ ID NO 11960.

RESULT 68
AAU32148
ID AAU32148 standard; protein; 592 AA.
XX
AC AAU32148;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #2639.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regenerati
KW immune suppression; immune stimulation; anti-inflammatory; leuk
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US008656.
XX
PR 18-APR-2000; 2000US-00552929.
PR 26-JAN-2001; 2001US-00770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in
PT vaccination, testing and therapy.
XX
PS Claim 20; Page 562-563; 765pp; English.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease assoc
CC with altered levels of polypeptide. The polypeptides are also us
CC identifying agents (agonists and antagonists) that bind to them.
CC expressing the proteins are useful for identifying a therapeutic
CC for use in treatment of a pathology related to aberrant expressi
CC physiological interactions of the polypeptide. Vectors comprisin
CC nucleic acids encoding the polypeptides and cells genetically en
CC to express them are also useful for producing the proteins. The
CC are useful in genetic vaccination, testing and therapy, and can
CC as nutritional supplements. They may be used to increase stem ce
CC proliferation; to regulate haematopoiesis; and in bone, cartilag
CC and/or nerve tissue growth or regeneration; immune suppression a
CC stimulation; as anti-inflammatory agents; and in treatment of le
CC AAU29510-AAU33304 represent the amino acid sequences of novel hu
CC secreted proteins of the invention
XX
SQ Sequence 592 AA;

Query Match 2.8%; Score 8; DB 4; Length 592;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; (

Qy 61 LGLGLALA 68
   ||||||||
Db 44 LGLGLALA 51

RESULT 69
ABG03722
ID ABG03722 standard; protein; 603 AA.
XX
AC ABG03722;
XX
DT 13-FEB-2002 (first entry)
XX

(first entry)

n P36269, SEQ ID NO 8914.

neuronal tissue; gene therapy;
ntal nerve injury; chronic construction injury; CCI;
; injury; SNI; Chung.

-A2.

2002WO-US025765.

2001US-0312147P.
2001US-0346382P.
2001US-0333347P.

HOSPITAL CORP.
R AG.

urso D, Befort K, Costigan M;

8312/26.
269.

ion comprising two or more isolated polypeptides, useful for
medicament for treating pain in an animal.

e; 1017pp; English.

n discloses a composition comprising two or more isolated rat
r allelic variation of the nucleic acid sequence. Also
a vector comprising the novel polynucleotide, a host cell
ne vector, a method for identifying a nucleotide sequence
ferentially regulated in an animal subjected to pain and a
rm the method, an array, a method for identifying an agent
es or decreases the expression of the polynucleotide sequence
erentially expressed in neuronal tissue of a first animal
pain, a method for identifying a compound which regulates
on of a polynucleotide sequence which is differentially
an animal subjected to pain, a method for identifying a
regulates the activity of one or more of the
des, a method for producing a pharmaceutical composition, a
dentifying a compound or small molecule that regulates the
an animal of one or more of the polypeptides given in the
1, a method for identifying a compound useful in treating
harmaceutical composition comprising the one or more
or their antibodies. The polynucleotide or the compound that
s activity is useful for preparing a medicament for treating
pinal segmental nerve injury (Chung), chronic constriction
and spared nerve injury (SNI) in an animal (e.g. gene
a sequence presented is a human protein (shown in Table 2 of
ntion) which is differentially expressed during pain. Note:
data for this patent did not form part of the printed
n, but was obtained in electronic form directly from WIPO at
/pub/published_pct_sequences.

AA;

larity 2.8%; Score 8; DB 7; Length 586;
Conservative 100.0%; Pred. No. 2.2e+02;
  0; Mismatches 0; Indels 0; Gaps 0;

GLALA 68
 |||
GLALA 21

KW　protein modification and maintenance molecule; PMMM;
KW　protein modification; protein maintenance; protein function;
KW　protein conformation; protein stabilisation; protein degradation;
KW　phosphatase; protease; protease inhibitor; isomerase; transferase
KW　molecular chaperone; anti-HIV; antiallergic; antiinflammatory;
KW　antianaemic; antiparkinsonian; nootropic; anticonvulsant;
KW　antiarteriosclerotic; antiasthmatic; immunosuppressive; antithyro
KW　cytostatic; hepatotropic; dermatological; antidiabetic; nephrotro
KW　antigout; thyromimetic; neuroprotective; osteopathic; antiarthrit
KW　antiparasitic; antihelmintic; antipsoriatic; uropathic; ophthalmo
KW　antirheumatic; haemostatic; antibacterial; virucide; protozoacide
KW　fungicide; gene therapy; cell proliferative disorder; arterioscle
KW　hepatitis; polycythaemia vera; psoriasis; primary thrombocytopae
KW　cancer; developmental disorder; anaemia; mental retardation;
KW　neurological disorder; Alzheimer's disease; Parkinson's disease;
KW　epilepsy; autoimmune disorder; inflammatory disorder; AIDS; aller
KW　asthma; autoimmune thyroiditis; Crohn's disease; diabetes mellitu
KW　glomerulonephritis; Goodpasture's syndrome; multiple sclerosis;
KW　arthritis; osteoporosis; pancreatitis; Sjogren's syndrome;
KW　microbial infection; human.
XX
OS　Homo sapiens.
XX
PN　WO200306368-A2.
XX
PD　07-AUG-2003.
XX
PF　23-JAN-2003; 2003WO-US002500.
XX
PR　25-JAN-2002; 2002US-0351928P.
PR　25-FEB-2002; 2002US-0359903P.
PR　21-MAR-2002; 2002US-0366837P.
XX
PA　(INCY-) INCYTE GENOMICS INC.
XX
PI　Hafalia AJA, Li JX, Gorvad AE, Chawla NK, Sprague WW, Lee SY,
PI　Chang H, Elliott VS, Ramkumar J, Khare R, Emerling BM, Kable
PI　Tang YT, Yue H, Gietzen KJ, Lee S, Swarnakar A, Baughn MR;
PI　Wilson AD, Jin P, Chien D, Hawkins PR, Jiang W, Jackson AA,
PI　Bhatia U, Burrill JD, Blake JJ, Ho A, Zheng W, Ison CH, Marc
PI　Tran UK, Lal PG, Warren BA, Xu Y, Honchell CD, Becha SD;
PI　Lehr-Mason PM;
XX
DR　WPI; 2003-636761/60.
DR　N-PSDB; ADE79064.
XX
PT　New human protein modification and maintenance molecules and
PT　polynucleotides, useful for diagnosing, treating or preventing aut
PT　or inflammatory disorders (e.g. AIDS, allergy or anemia), multiple
PT　sclerosis or cancer.
XX
PS　Claim 1; SEQ ID NO 44; 405pp; English.
XX
CC　This invention relates to novel isolated human proteins, which are
CC　protein modification and maintenance molecules (PMMM). The cellula
CC　processes regulating modification and maintenance of protein molec
CC　coordinate their function, conformation, stabilisation and degrada
CC　Each of these processes is mediated by key enzymes or proteins suc
CC　kinases, phosphatases, proteases, protease inhibitors, isomerases,
CC　transferases and molecular chaperones. Compounds which modulate th
CC　proteins of the invention may have anti-HIV, antiallergic,
CC　antiinflammatory, antianaemic, antiparkinsonian, nootropic,
CC　anticonvulsant, antiarteriosclerotic, antiasthmatic, immunosuppres
CC　antithyroid, cytostatic, hepatotropic, dermatological, antidiabeti
CC　nephrotropic, antigout, thyromimetic, neuroprotective, osteopathic
CC　antiarthritic, antiparasitic, antihelmintic, antipsoriatic, uropat
CC　ophthalmological, antirheumatic, haemostatic, antibacterial, viruc
CC　protozoacide or fungicide activities. The DNA sequence which encod
CC　proteins of the invention may be useful for gene therapy. The huma
CC　protein modification and maintenance molecules (PMMB), the DNA se
CC　which encode them and their modulating compounds are useful for
CC　diagnosing, treating or preventing disorders associated with aberr
CC　expression of PMMM, particularly cell proliferative disorders (for

---

agnostic protein #3713.

some mapping; gene mapping; gene therapy; forensic;
t; medical imaging; diagnostic; genetic disorder.

2001WO-US008631.

2000US-00540217.
2000US-00649167.

INC.

iu C, Tang YT;

62/73.
09.

olynucleotide and encoded polypeptides, useful in
orensics, gene mapping, identification of mutations
r genetic disorders or other traits and to assess

ID NO 34081; 103pp; English.

relates to isolated polynucleotide (I) and polypeptide (II)
is useful as hybridisation probes, polymerase chain
primers, oligomers, and for chromosome and gene mapping,
nant production of (II). The polynucleotides are also used
as expressed sequence tags for identifying expressed
useful in gene therapy techniques to restore normal
I) or to treat disease states involving (II). (II) is
erating antibodies against it, detecting or quantitating a
tissue, as molecular weight markers and as a food
I) and its binding partners are useful in medical imaging
ssing (II). (I) and (II) are useful for treating disorders
rant protein expression or biological activity. The
d polynucleotide sequences have application in
orensics, gene mapping, identification of mutations
r genetic disorders or other traits to assess biodiversity
other types of data and products dependent on DNA and
uences. ABG00010-ABG30377 represent novel human diagnostic
uences of the invention. Note: The sequence data for this
appear in the printed specification, but was obtained in
mat directly from WIPO at
ub/published_pct_sequences

A;

rity　　2.8%;　Score 8;　DB 4;　Length 603;
rity　100.0%;　Pred. No. 2.3e+02;　Indels　　0;　Gaps　　0;
nservative　0;　Mismatches　0;

ALA 68
|||
ALA 447

ard; protein; 617 AA.

first entry)

nodification and maintenance molecule (PMMM)-44.

CC   diagnostics, forensics, gene mapping, identification of mutation
CC   responsible for genetic disorders or other traits to assess bioc
CC   and to produce other types of data and products dependent on DNA
CC   amino acid sequences. ABG0010-ABG30377 represent novel human di
CC   amino acid sequences of the invention. Note: The sequence data f
CC   patent did not appear in the printed specification, but was obta
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 633 AA;

Query Match   2.8%; Score 8; DB 4; Length 633;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches   8; Conservative   0; Mismatches   0; Indels   0;

Qy   61 LGLGLALA 68
                |||||||
Db   44 LGLGLALA 51

RESULT 72
ABG28291
ID   ABG28291 standard; protein; 799 AA.
XX
AC   ABG28291;
XX
DT   18-FEB-2002 (first entry)
XX
DE   Novel human diagnostic protein #28282.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT, Liu C, Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS92478.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutation;
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 20; SEQ ID NO 58650; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypept
CC   sequences. (I) is useful as hybridisation probes, polymerase chai
CC   reaction (PCR) primers, oligomers, and for chromosome and gene ma
CC   and in recombinant production of (II). The polynucleotides are al
CC   in diagnostics as expressed sequence tags for identifying express
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II)
CC   useful for generating antibodies against it, detecting or quantit
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical i
CC   of sites expressing (II). (I) and (II) are useful for treating di
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodi

Query Match 2.8%; Score 8; DB 7; Length 842;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; C

```
Qy    54 ALLVPLAL 61
         ||||||||
Db   441 ALLVPLAL 448
```

RESULT 74
ABU16705
ID   ABU16705 standard; protein; 1032 AA.
XX
AC   ABU16705;
XX
DT   19-JUN-2003 (first entry)
XX
DE   Protein encoded by Prokaryotic essential gene #2232.
XX
KW   Antisense; prokaryotic essential gene; cell proliferation; drug d
XX
OS   Acinetobacter baumannii.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind
PI   Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   N-PSDB; ACA20575.
XX
PT   New antisense nucleic acids, useful for identifying proteins or s
PT   for homologous nucleic acids required for cellular proliferation
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 25; SEQ ID NO 44629; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any
CC   the 6213 antisense sequences given in the specification where exp
CC   of the nucleic acid inhibits proliferation of a cell. Also include
CC   (1) a vector comprising a promoter operably linked to the nucleic
CC   encoding a polypeptide whose expression is inhibited by the antise
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolat
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically bi
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cell
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the acti
CC   the gene product or that has an activity against a biological path
CC   required for proliferation, or that inhibits cellular proliferatio
CC   identifying a gene required for cellular proliferation or the biol
CC   pathway in which a proliferation-required gene or its gene product
CC   or a gene on which the test compound that inhibits proliferation c
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which th
CC   product is overexpressed or underexpressed; (12) determining the e
CC   to which each of the strains is present in a culture or collection
CC   strains; or (13) identifying the target of a compound that inhibit
CC   proliferation of an organism. The antisense nucleic acids are usef
CC   identifying proteins or screening for homologous nucleic acids req

---

e. other types of data and products dependent on DNA and
quences. ABG00010-ABG30377 represent novel human diagnostic
quences of the invention. Note: The sequence data for this
t appear in the printed specification, but was obtained in
pub/published_pct_sequences

AA;

arity 100.0%; Pred. No. 3e+02;
onservative 0; Mismatches 0; Indels 0; Gaps 0;

```
GLAL 67
||||
GLAL 294
```

dard; protein; 842 AA.

(first entry)

(useful for identifying genetic disorders) #630.

ovel protein; tissue marker; molecular weight marker;
ker; genetic disorder.

-2.

002WO-US039555.

001US-0339739P.
001US-0339453P.
002US-0365091P.
002US-0365384P.
002US-0372381P.
002US-0372615P.
002US-00128558.
002US-0376045P.

INC.

di V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;
Chen R, Xu C, Boyle BJ;

35/53.
64.

tides, useful for expressing recombinant proteins for
acterization or therapeutic use, or as markers for tissues
orresponding protein is preferentially expressed.

ID NO 1541; 1177pp; English.

comprises the amino acid and coding sequences of novel
DNA and protein sequences of the invention are useful as:
ssues in which the corresponding protein is preferentially
molecular weight markers on gels; as chromosome markers or
ify chromosomes or to map related gene positions; and to
ndogenous DNA sequences in patients to identify potential
ers. The present amino acid sequence represents a protein
on.

A;

proliferation to isolate candidate molecules for rational
ry programs, or for screening homologous nucleic acids
proliferation in cells other than S. aureus, S. typhimurium,
e or P. aeruginosa. The present sequence is encoded by one of
rokaryotic essential genes. Note: The sequence data for this
ot form part of the printed specification, but was obtained
c format directly from WIPO at
/pub/published_pct_sequences

2 AA;

larity   2.8%;   Score 8; DB 6; Length 1032;
Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

ALGLG 64
|||||
ALGLG 981

ndard; protein; 1033 AA.

(first entry)

r baumannii protein #1012.

r baumannii; bacterial disease; antibacterial; vaccine;
rol agent.

r baumannii.

99US-00328352.

98US-0088701P.

4E THERAPEUTICS CORP.

ish D;

5092/54.
9725.

cter baumannii proteins and nucleic acids, useful as reagents
g a bacterial disease, as components of antibacterial
targets for antibacterial drugs, or as biocontrol agents for

ID NO 5138; 328pp; English.

1 relates to isolated Acinetobacter baumannii nucleic acids.
nii nucleic acids and polypeptides are useful as reagents
g a bacterial disease, as components of antibacterial
targets for antibacterial drugs, to detect the presence of
and other Acinetobacter species in a sample, in screening
the ability to interfere with the A. baumannii life cycle
A. baumannii infection, and as biocontrol agents for
present sequence represents the amino acid sequence of an A.
otein.

AA;

arity   2.8%;   Score 8; DB 6; Length 1033;
onservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   57 VPLALGLG 64
             ||||||||
Db   975 VPLALGLG 982

Search completed: April 7, 2004, 17:57:27
Job time : 65 secs